



US009070192B1

(12) **United States Patent**
Smith

(10) **Patent No.:** **US 9,070,192 B1**
(45) **Date of Patent:** **Jun. 30, 2015**

(54) **IMPLEMENTING RICH COLOR
TRANSITION CURVE TRACKING FOR
APPLICATIONS**

(71) Applicant: **VISION INTERFACE
TECHNOLOGIES, LLC**, Ann Arbor,
MI (US)

(72) Inventor: **Peter Harmon Smith**, Ann Arbor, MI
(US)

(73) Assignee: **VISION INTERFACE
TECHNOLOGIES, LLC**, Ann Arbor,
MI (US)

(*) Notice: Subject to any disclaimer, the term of this
patent is extended or adjusted under 35
U.S.C. 154(b) by 33 days.

(21) Appl. No.: **14/014,936**

(22) Filed: **Aug. 30, 2013**

Related U.S. Application Data

(63) Continuation-in-part of application No. 13/301,216,
filed on Nov. 21, 2011, now Pat. No. 8,526,717, which
is a continuation-in-part of application No.
12/107,092, filed on Apr. 22, 2008, now Pat. No.
8,064,691.

(60) Provisional application No. 60/917,966, filed on May
15, 2007.

(51) **Int. Cl.**
G06K 9/00 (2006.01)
G06T 7/00 (2006.01)

(52) **U.S. Cl.**
CPC **G06T 7/0042** (2013.01)

(58) **Field of Classification Search**
CPC G06T 5/002; G06T 7/0042; G06T 7/602;
G06T 7/606; H04N 5/217; H04N 1/54
USPC 382/133, 162, 164, 167, 172, 190, 224,
382/260, 261, 262, 264, 276, 277
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,795,068 B1 *	9/2004	Marks	345/419
2008/0056607 A1 *	3/2008	Ovsiannikov	382/275
2010/0172567 A1 *	7/2010	Prokoski	382/132

* cited by examiner

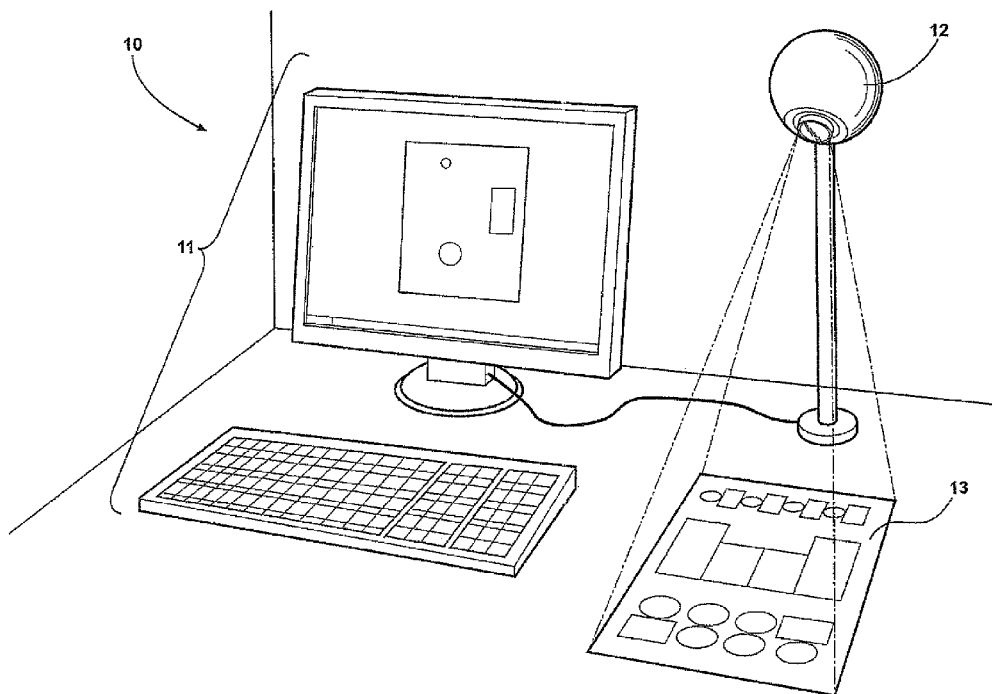
Primary Examiner — Ali Bayat

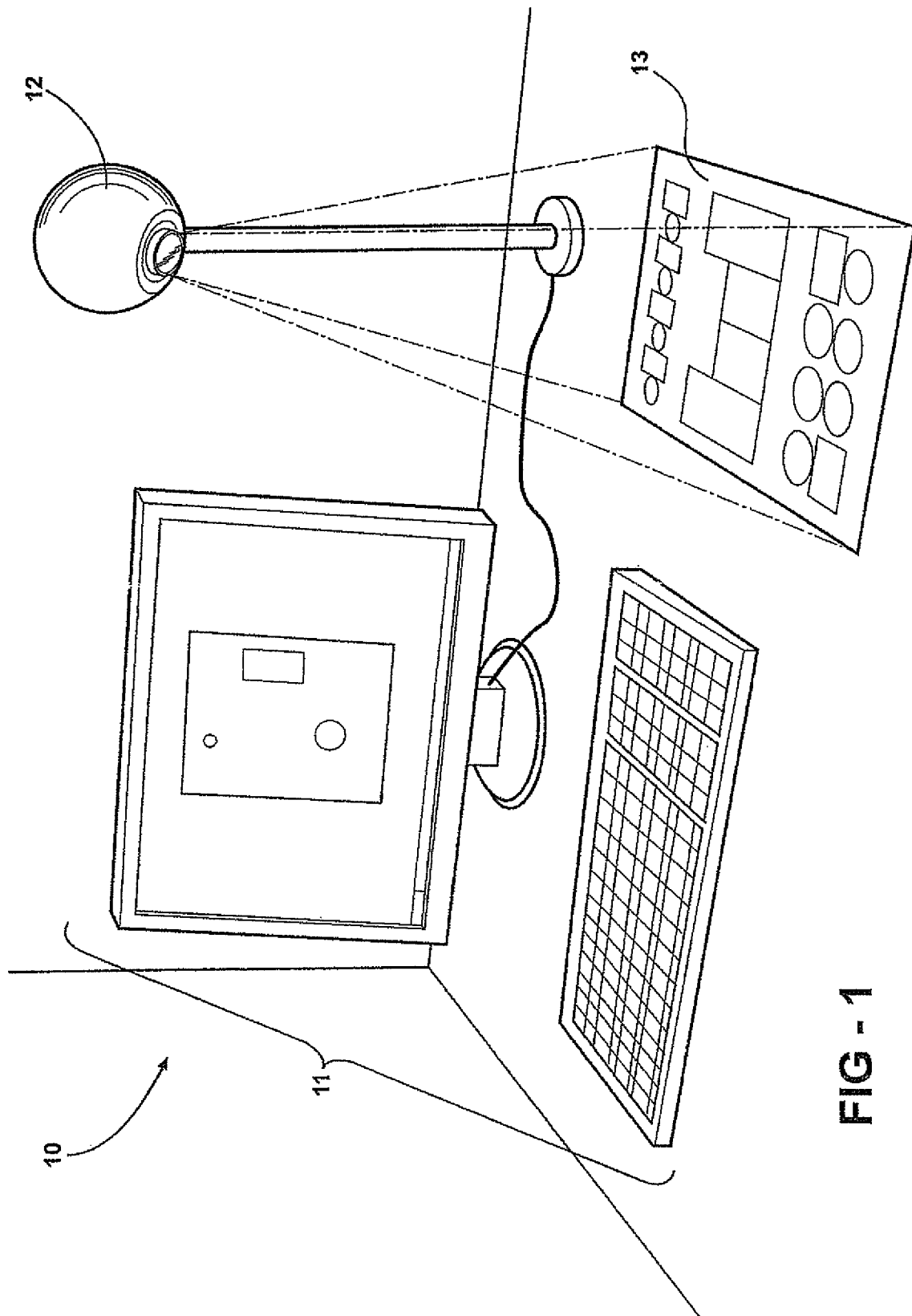
(74) *Attorney, Agent, or Firm* — Young Basile Hanlon &
MacFarlane P.C.

(57) **ABSTRACT**

A method of finding unique targets arranged in a unique
sequence of targets in a color image applies ratio space rich
colored filtering to identify a band of pixels that form a
transition curve from one rich color to another. A set of
transition curves with simple geometric and color properties
for the entire image, are compared to a database of target
properties to identify a target in an image. The targets are
provided on a carrier that is attached to a surface to enable
processor to identify the surface location and orientation by
the targets.

20 Claims, 34 Drawing Sheets





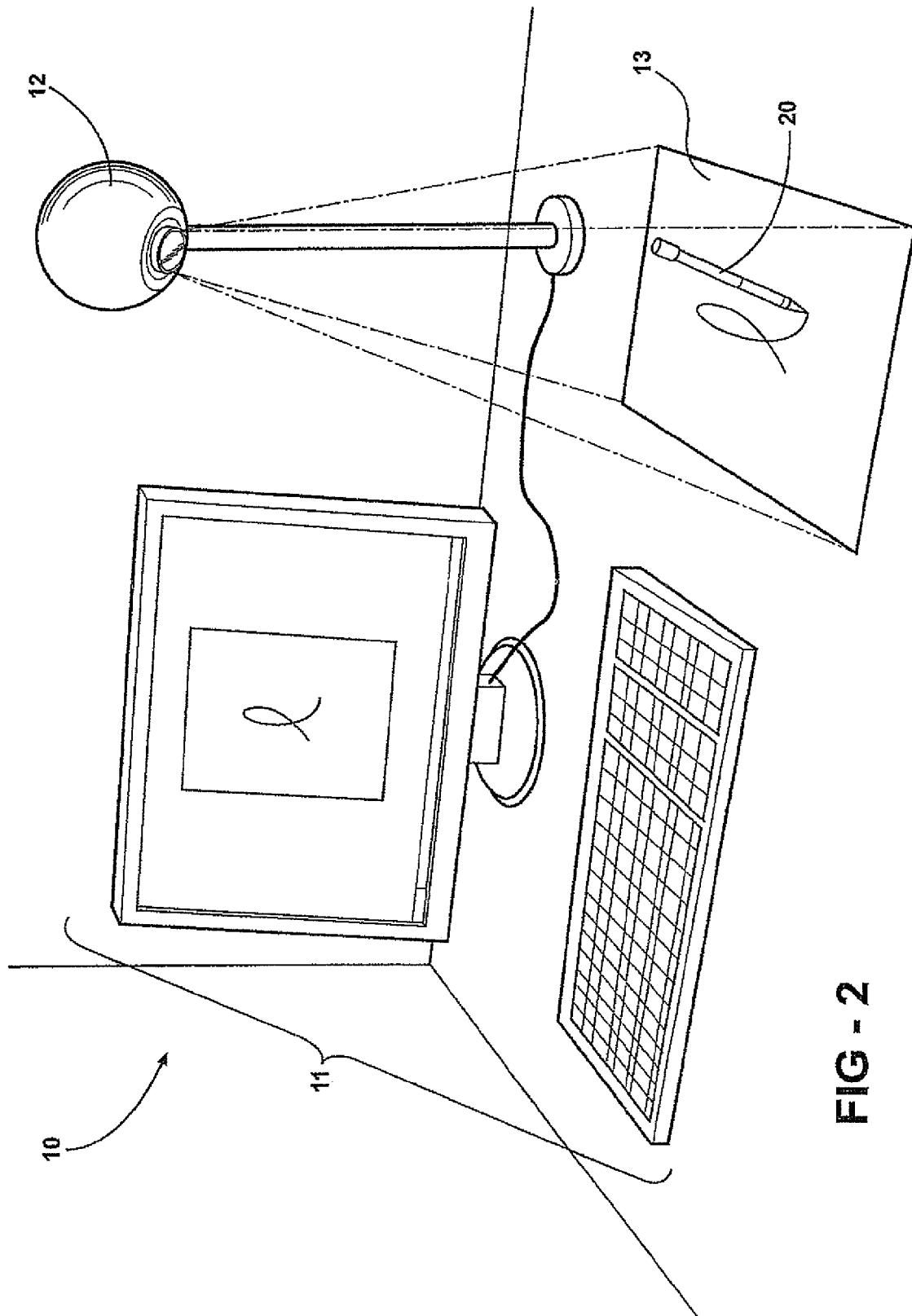


FIG - 2

Figure 3

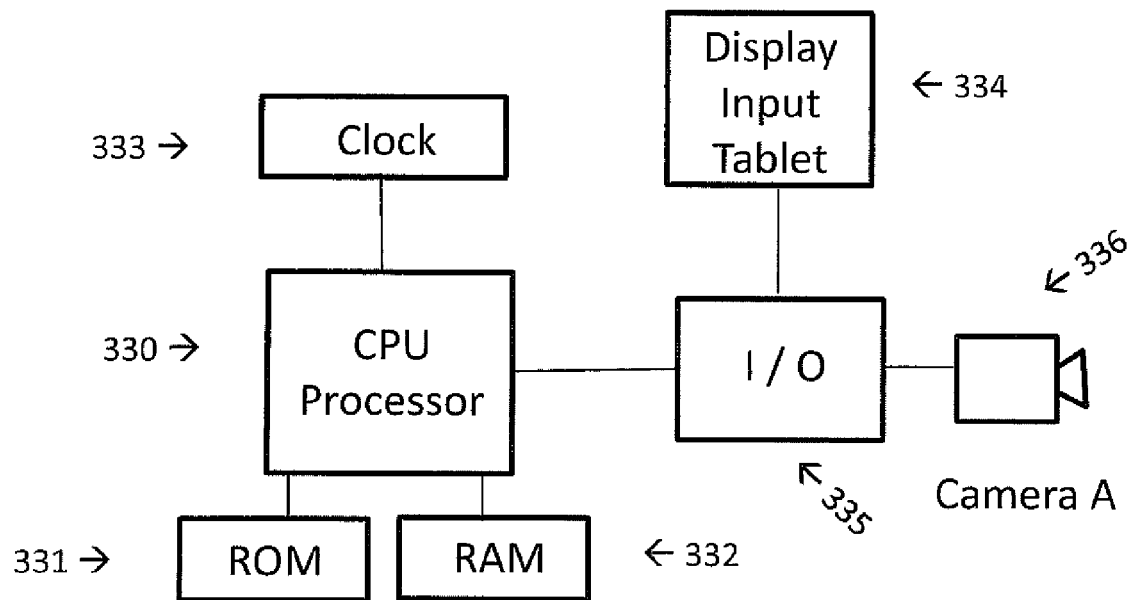


Figure 4

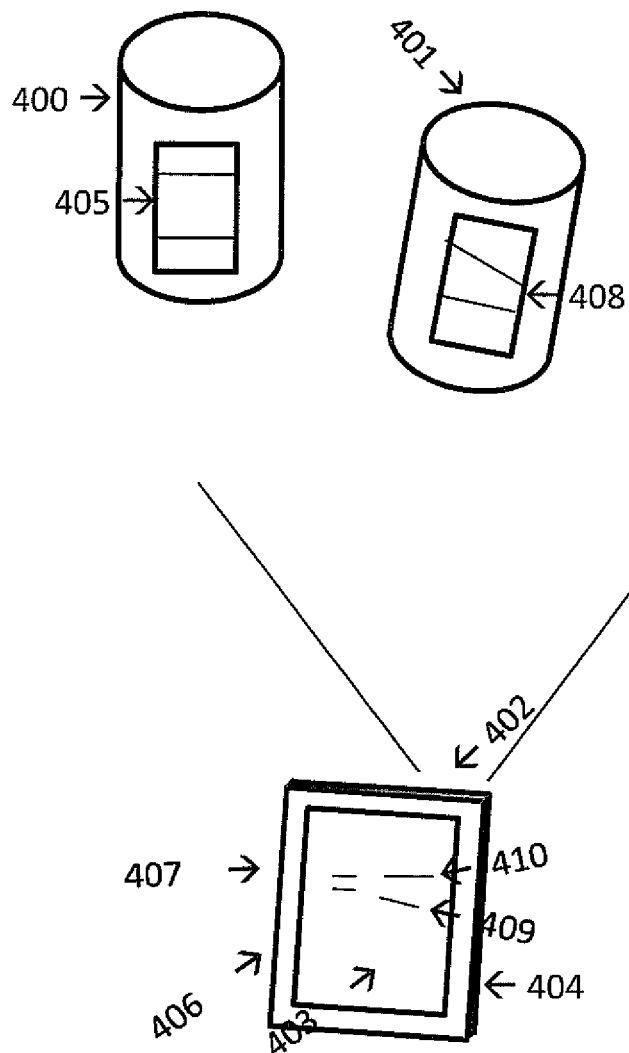


Figure 5

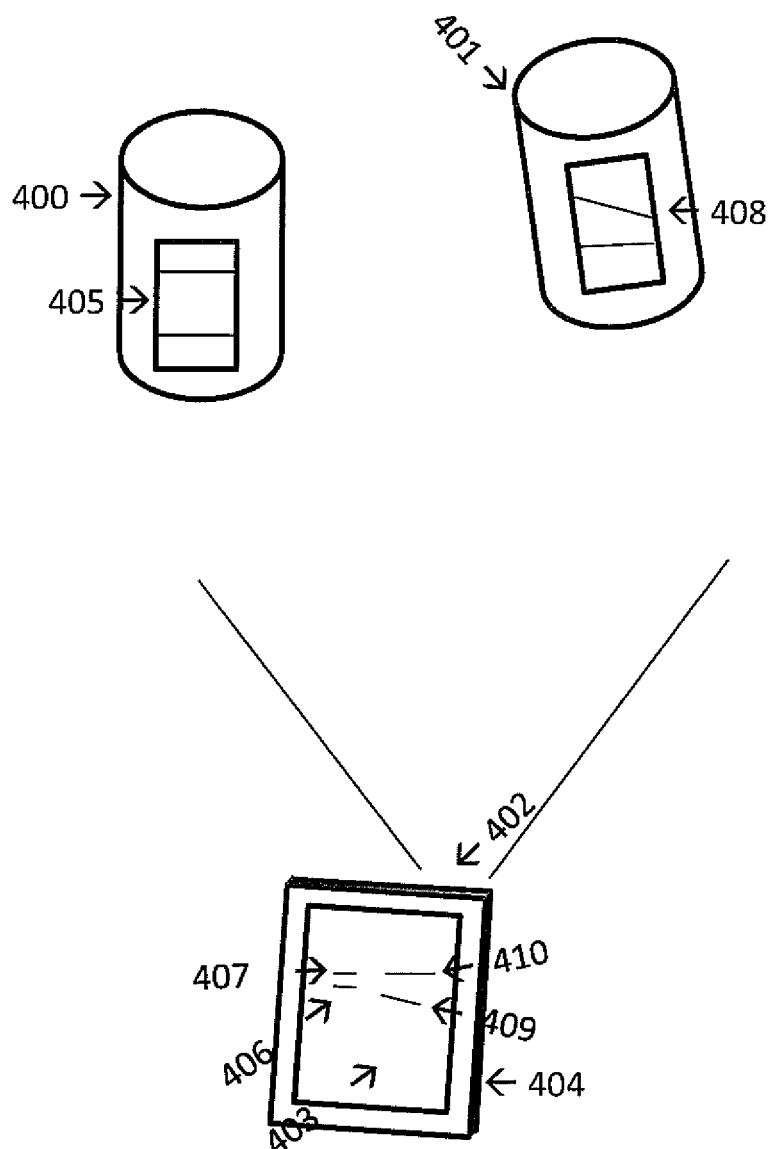


Figure 6

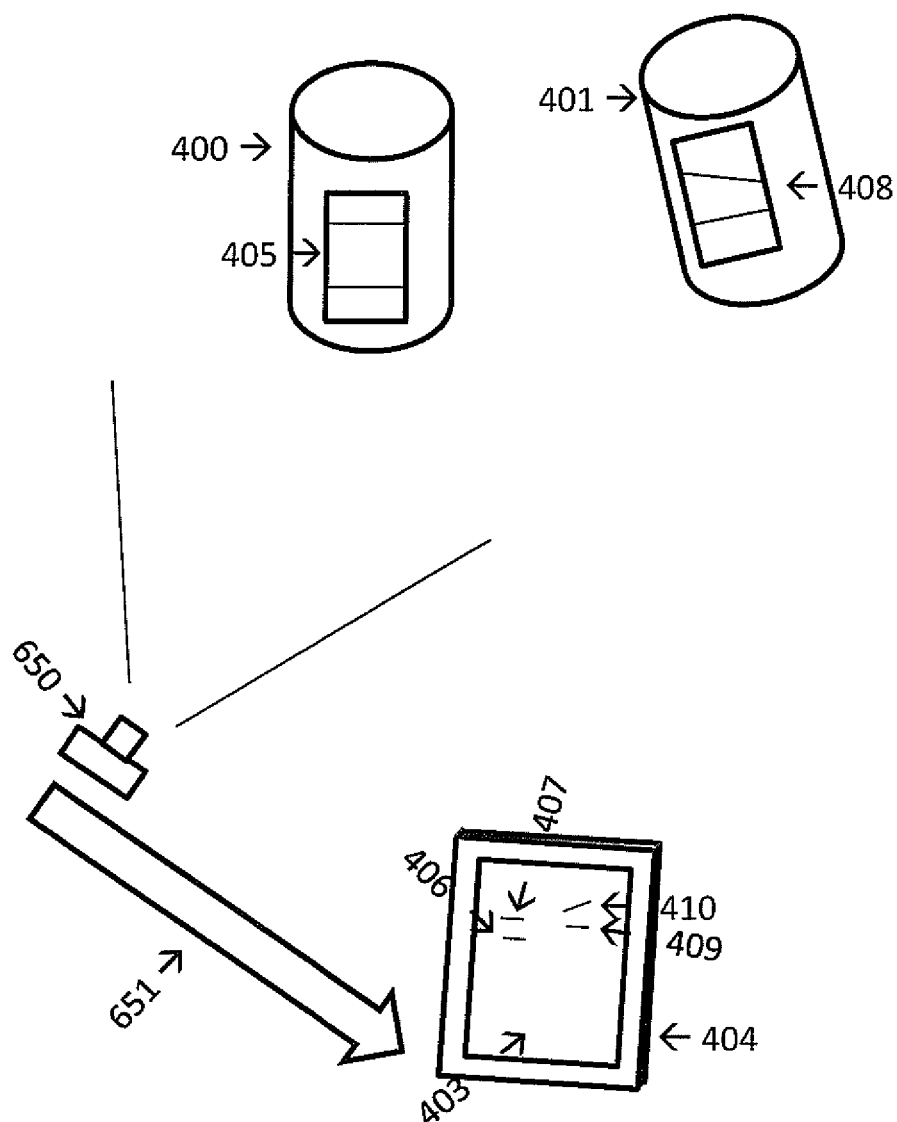


Figure 7

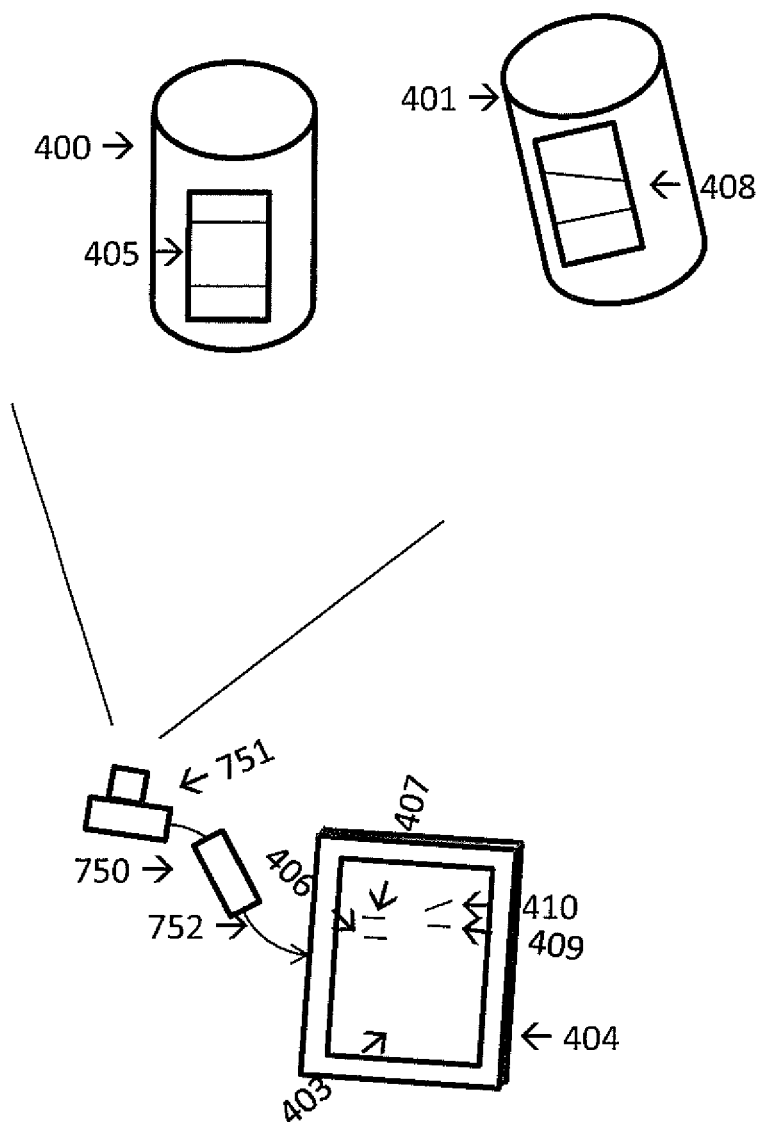


Figure 8

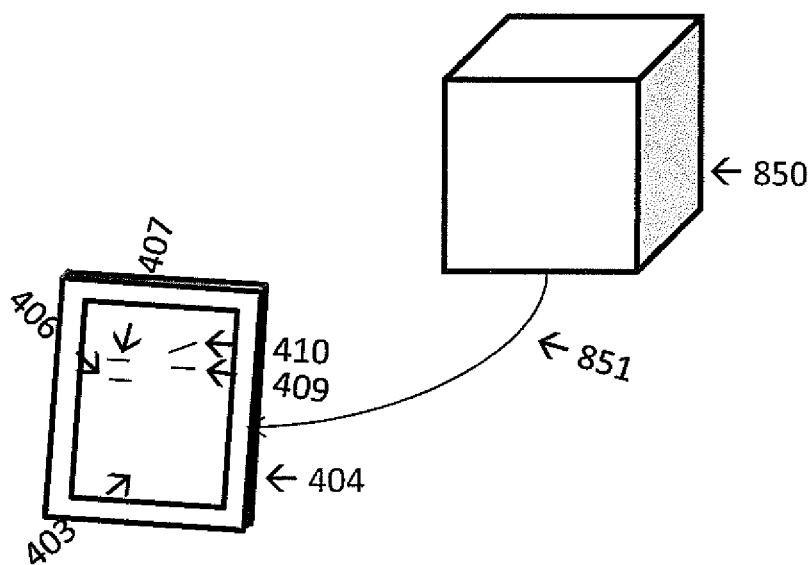


Figure 9

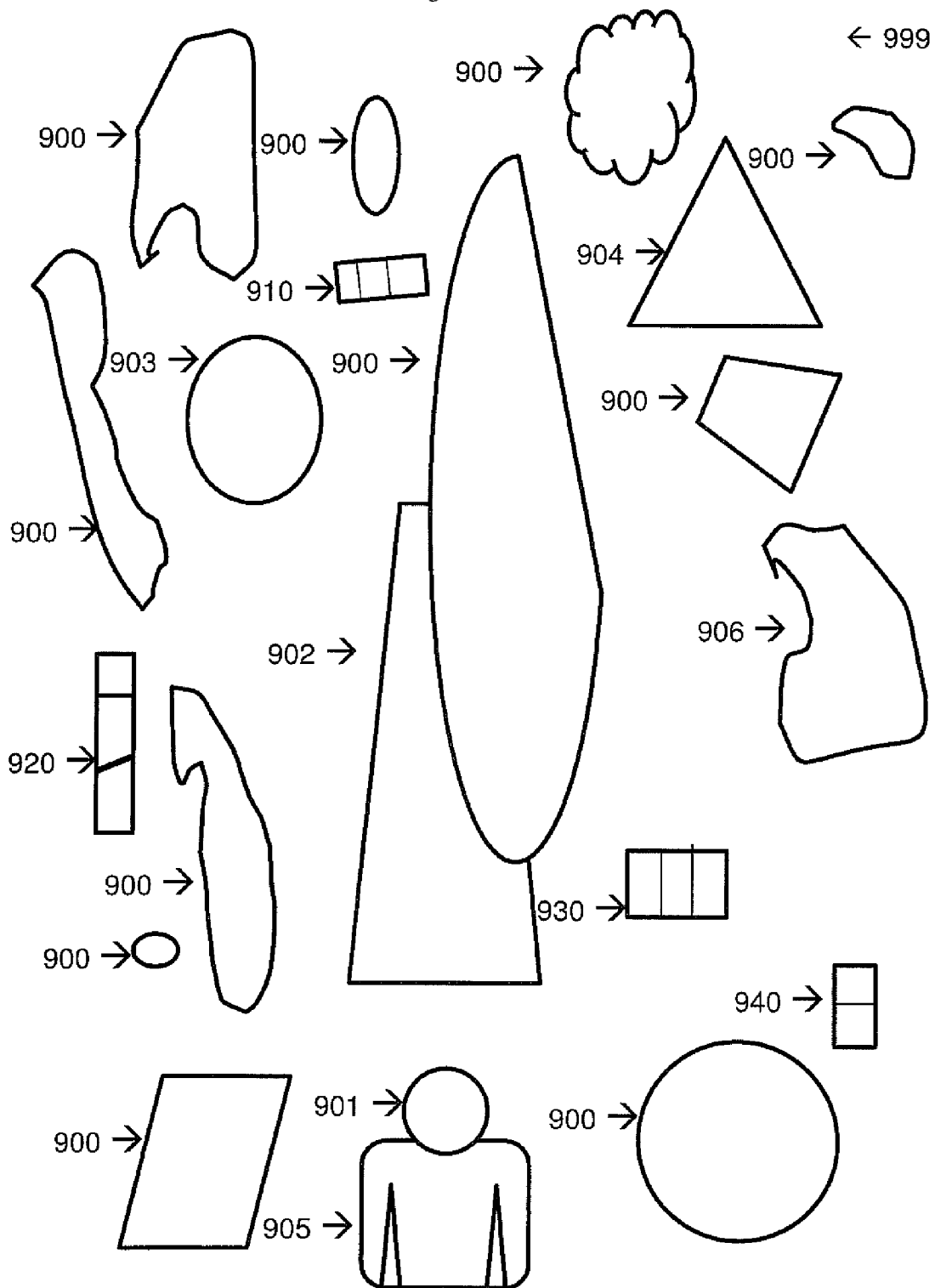


Figure 10

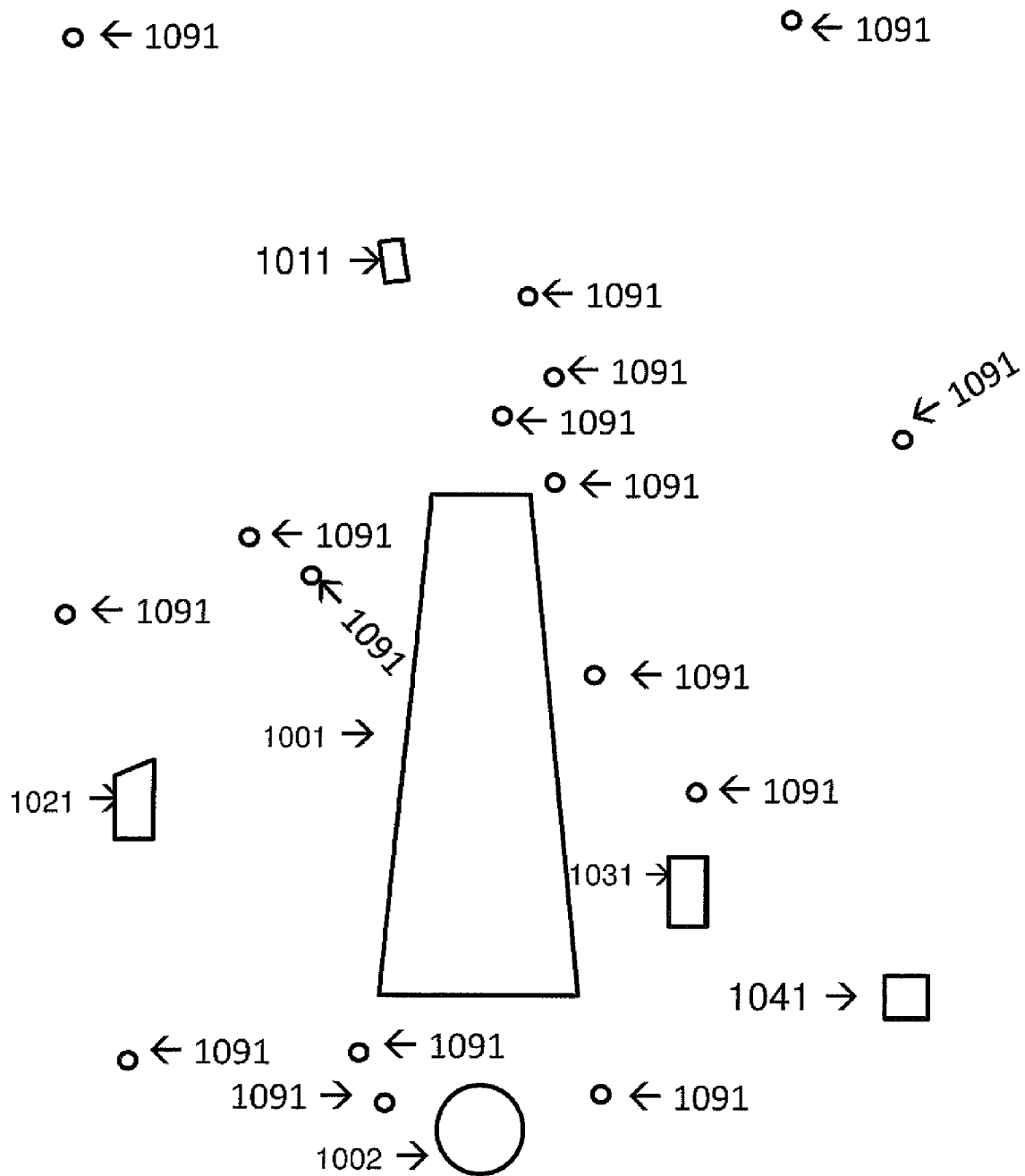


Figure 11

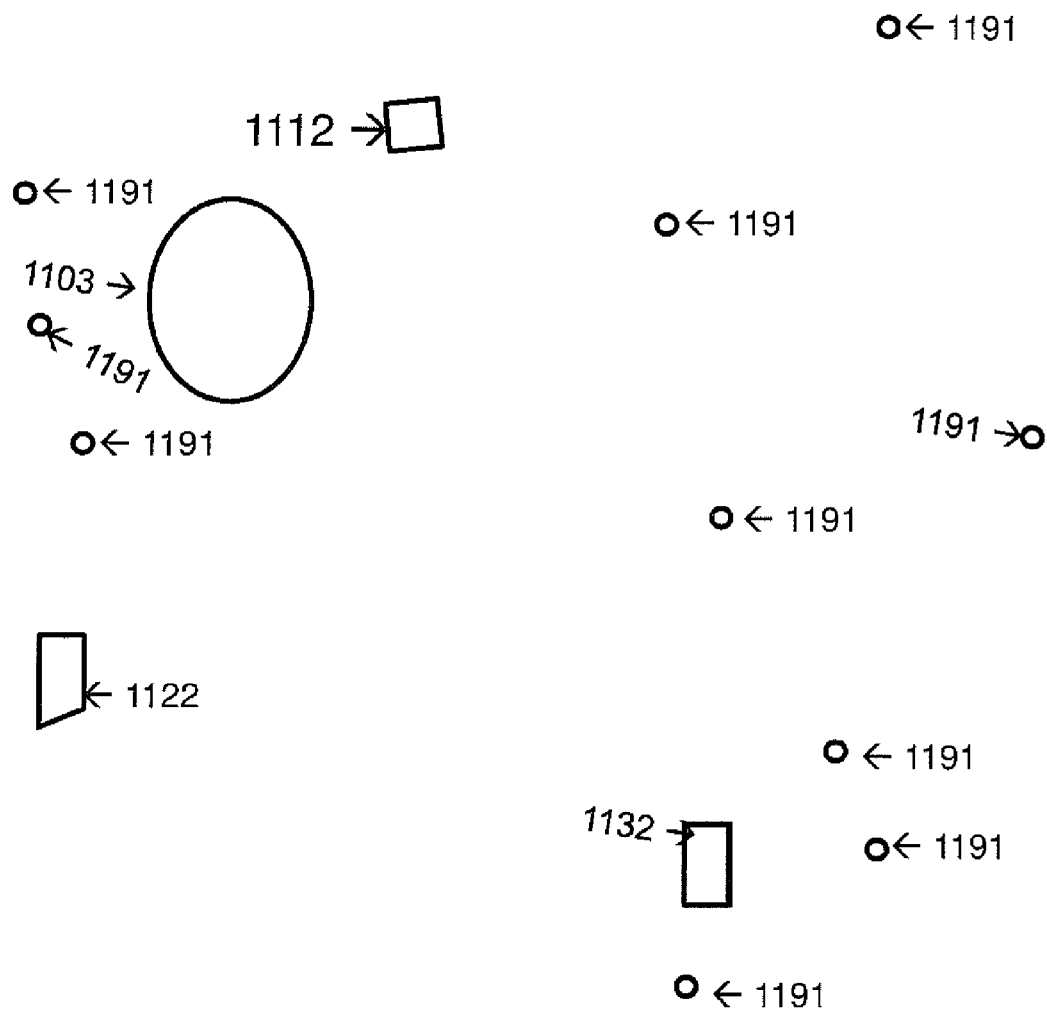


Figure 12

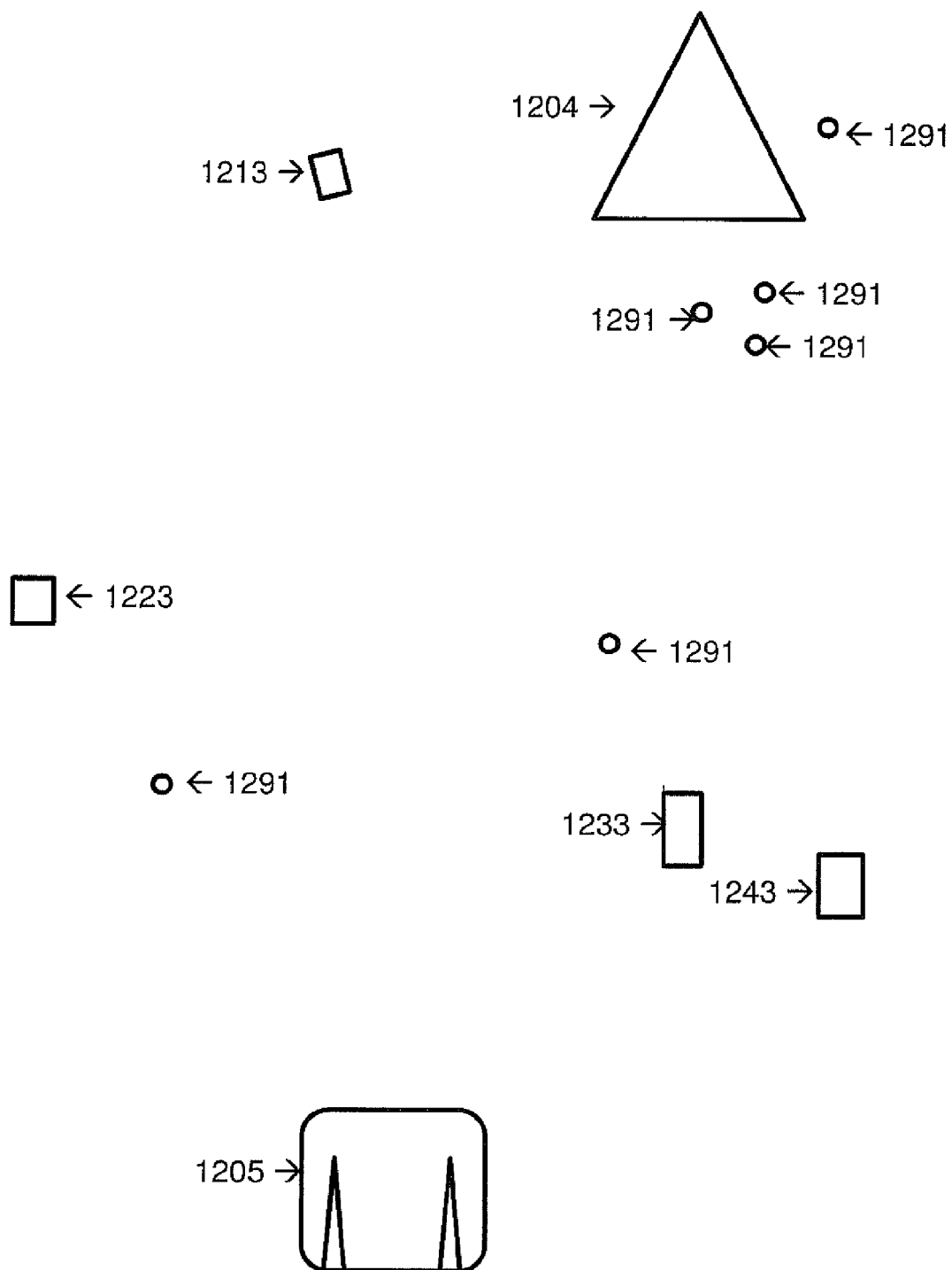


Figure 13

○← 1391

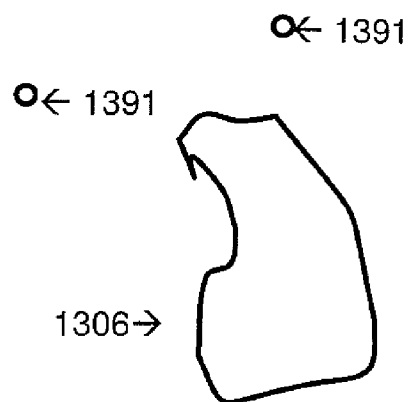


Figure 14

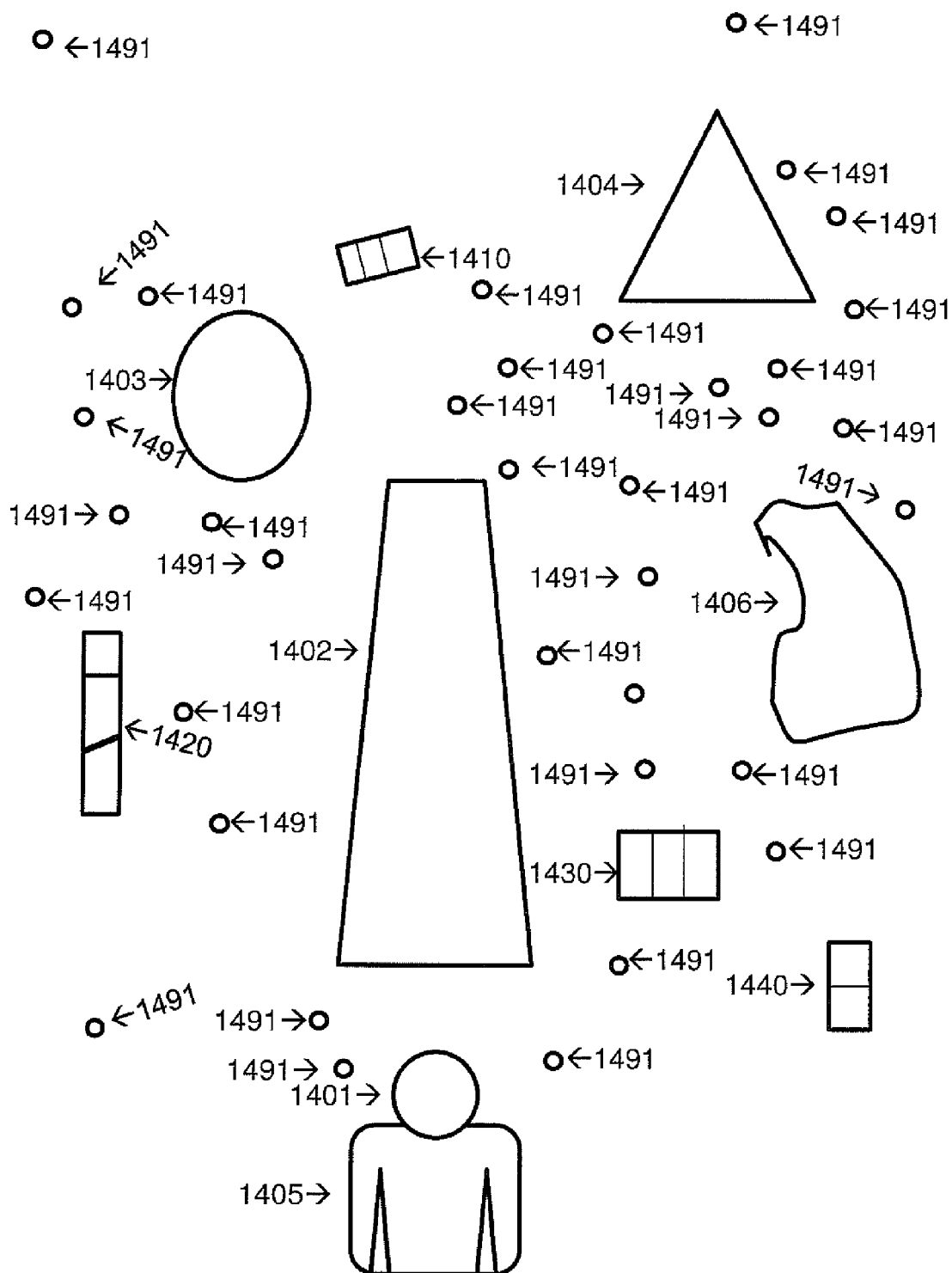


Figure 15

←1525

←1535

Figure 16

\ ←1617

1647→ —

←1657

Figure 17

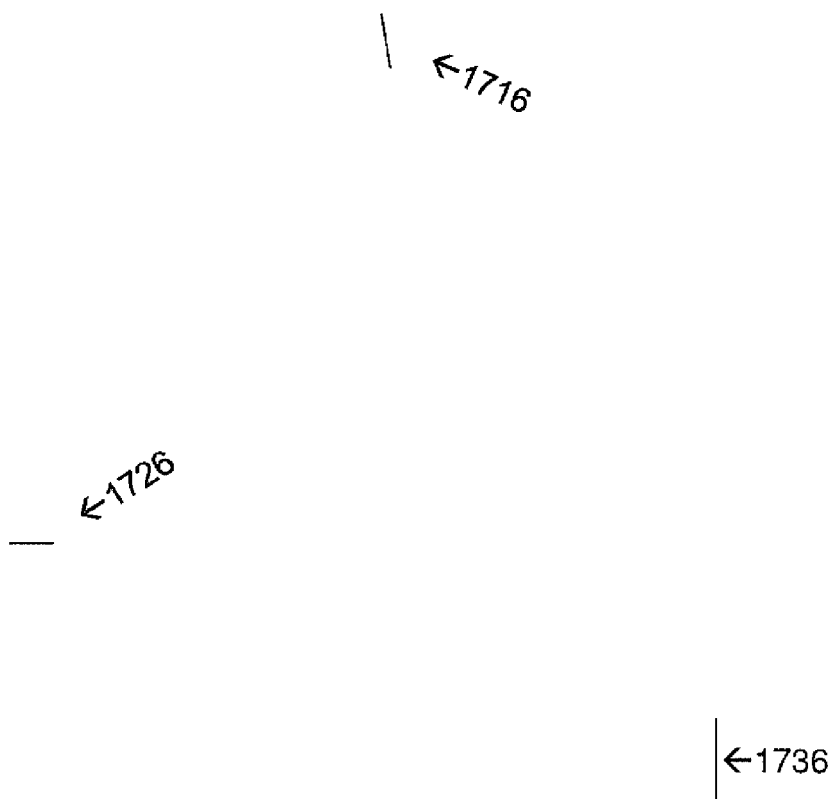


Figure 18

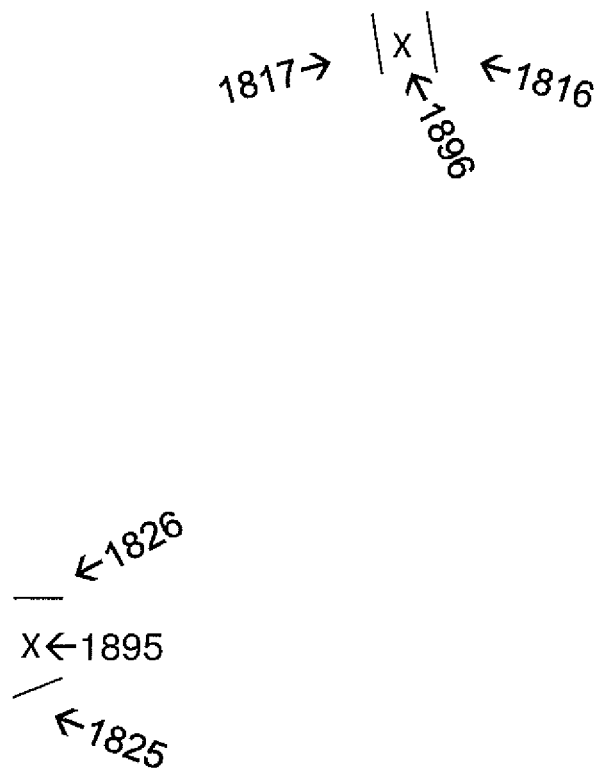


Figure 19

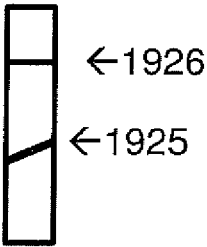
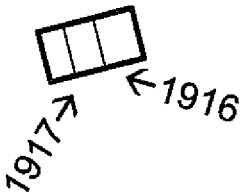


Figure 20

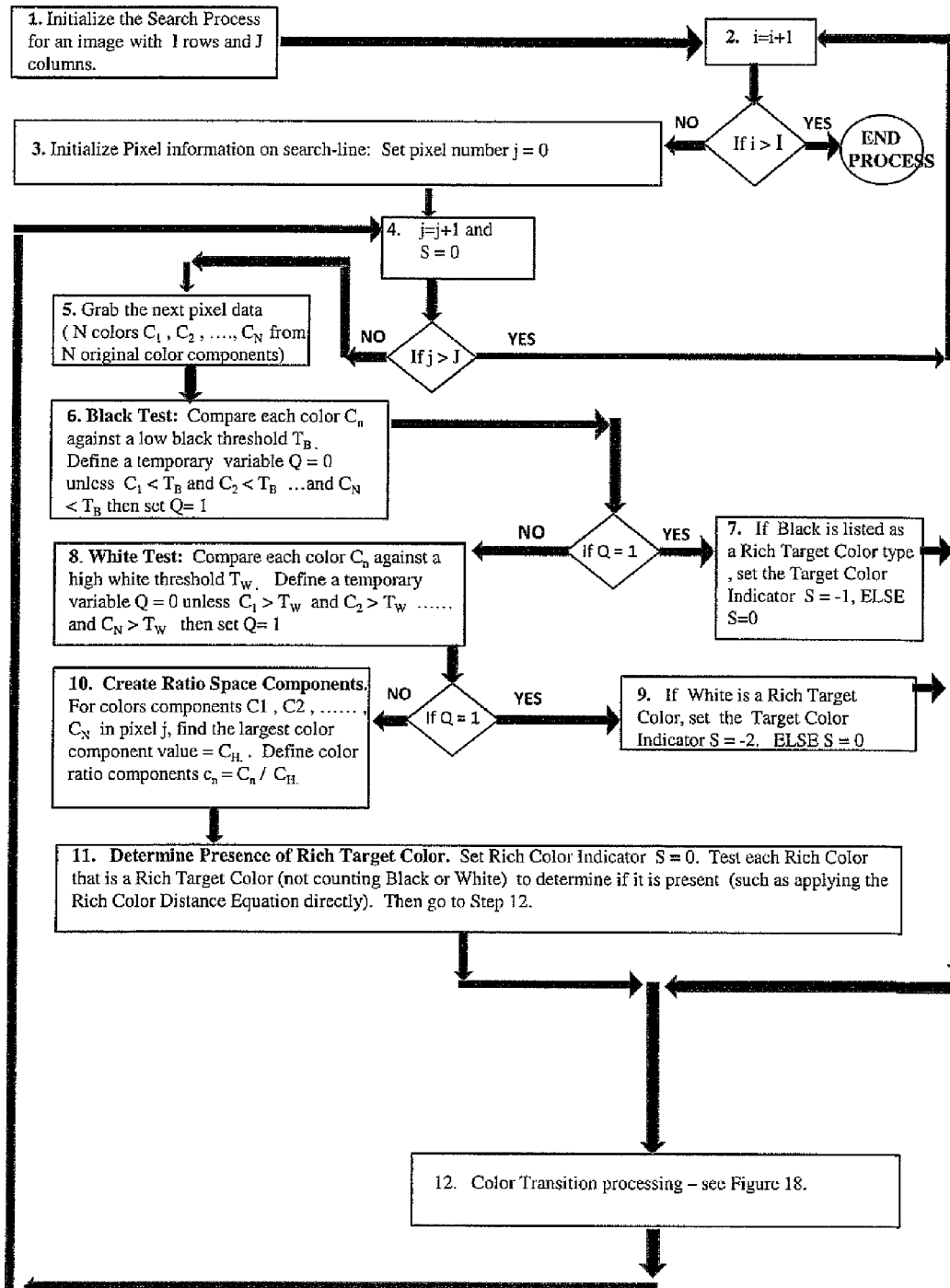


Figure 21

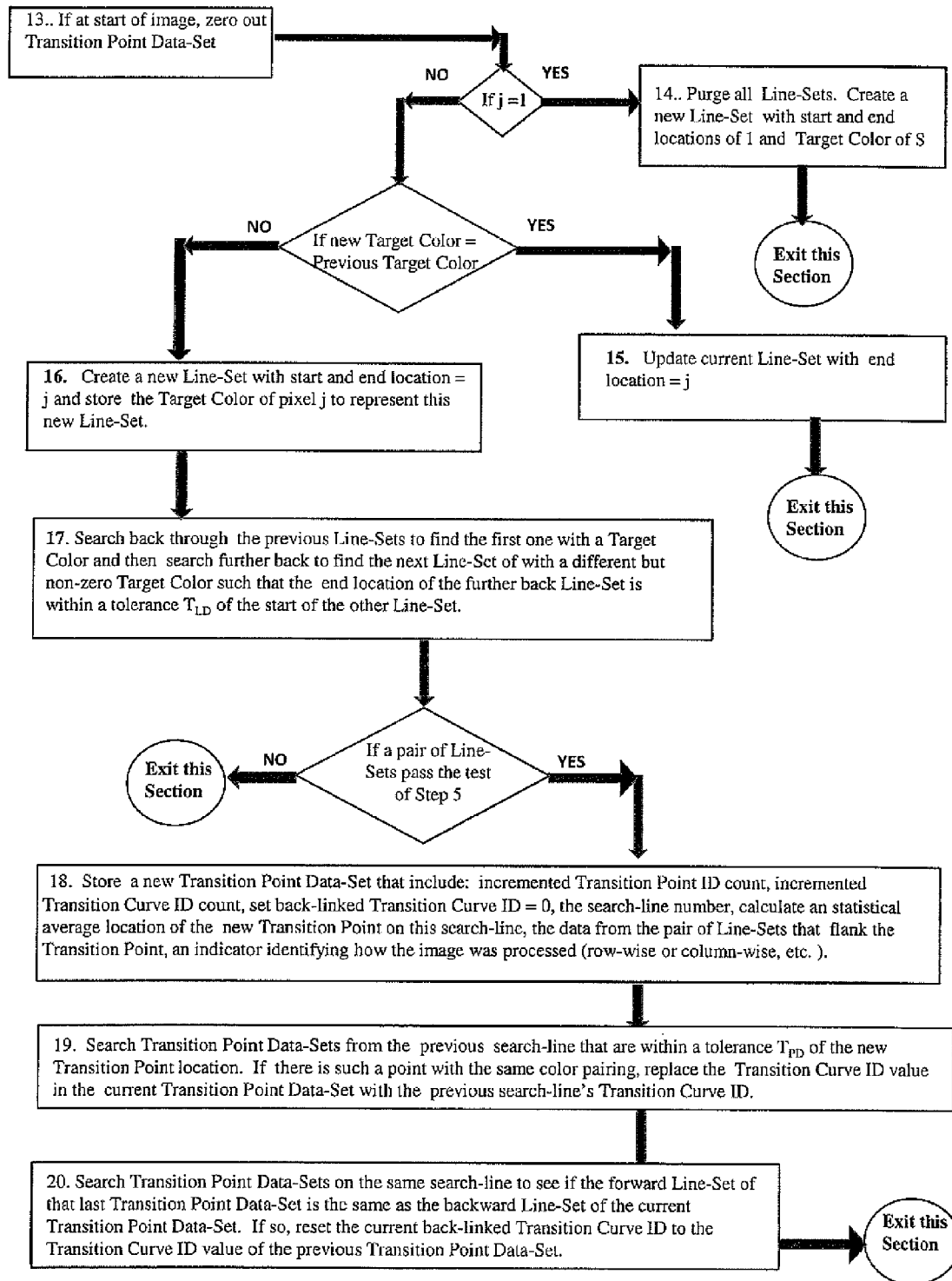


Figure 22

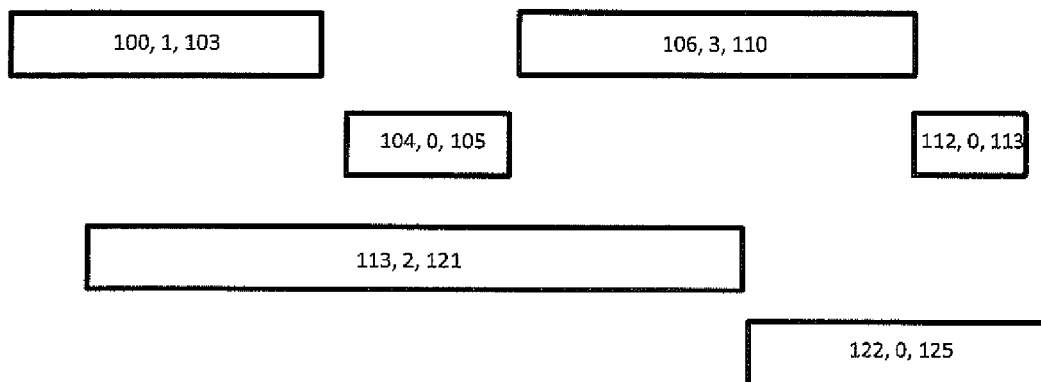
J = 100 101 102 103 104 105 106 107 108 109 110 111 112

RED S=1	RED S=1	RED S=1	RED S=1	Non Target Color	Non Target Color	Blue S=3	Blue S=3	Blue S=3	Blue S=3	Blue S=3	Non Target Color	Non Target Color
------------	------------	------------	------------	------------------------	------------------------	-------------	-------------	-------------	-------------	-------------	------------------------	------------------------

J = 113 114 115 116 117 118 119 120 121 122 123 124 125

Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Non Target Color	Non Target Color	Non Target Color	Non Target Color
--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	------------------------	------------------------	------------------------	------------------------

Line-Sets within Separation Tolerance with different Target Colors



Transition Point Data-Sets for Separation Tolerance = 4
and Target Colors of Red, Green, Blue

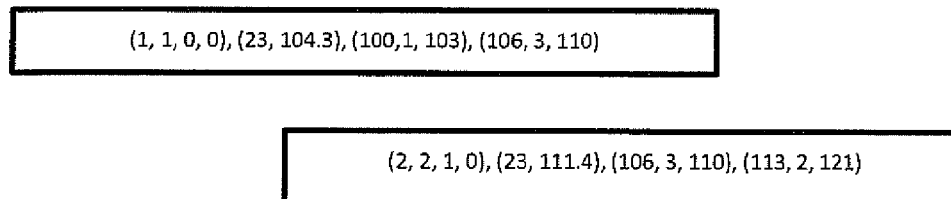


Figure 23

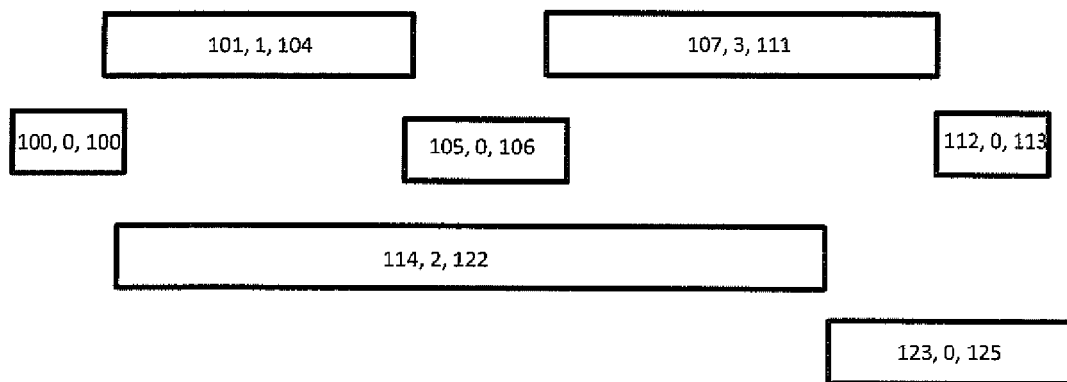
J = 100 101 102 103 104 105 106 107 108 109 110 111 112

Non Target Color	RED S=1	RED S=1	RED S=1	RED S=1	Non Target Color	Non Target Color	Blue S=3	Blue S=3	Blue S=3	Blue S=3	Blue S=3	Non Target Color
------------------------	------------	------------	------------	------------	------------------------	------------------------	-------------	-------------	-------------	-------------	-------------	------------------------

J = 113 114 115 116 117 118 119 120 121 122 123 124 125

Non Target Color	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Green S=2	Non Target Color	Non Target Color	Non Target Color
------------------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	------------------------	------------------------	------------------------

Line-Sets within Separation Tolerance with different Target Colors



Transition Point Data-Sets for Separation Tolerance = 4
and Target Colors or Red, Green, Blue

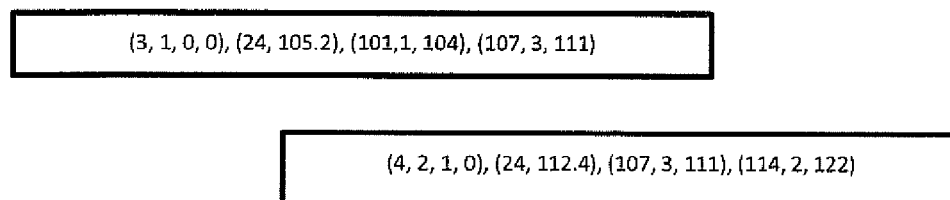


FIG 24

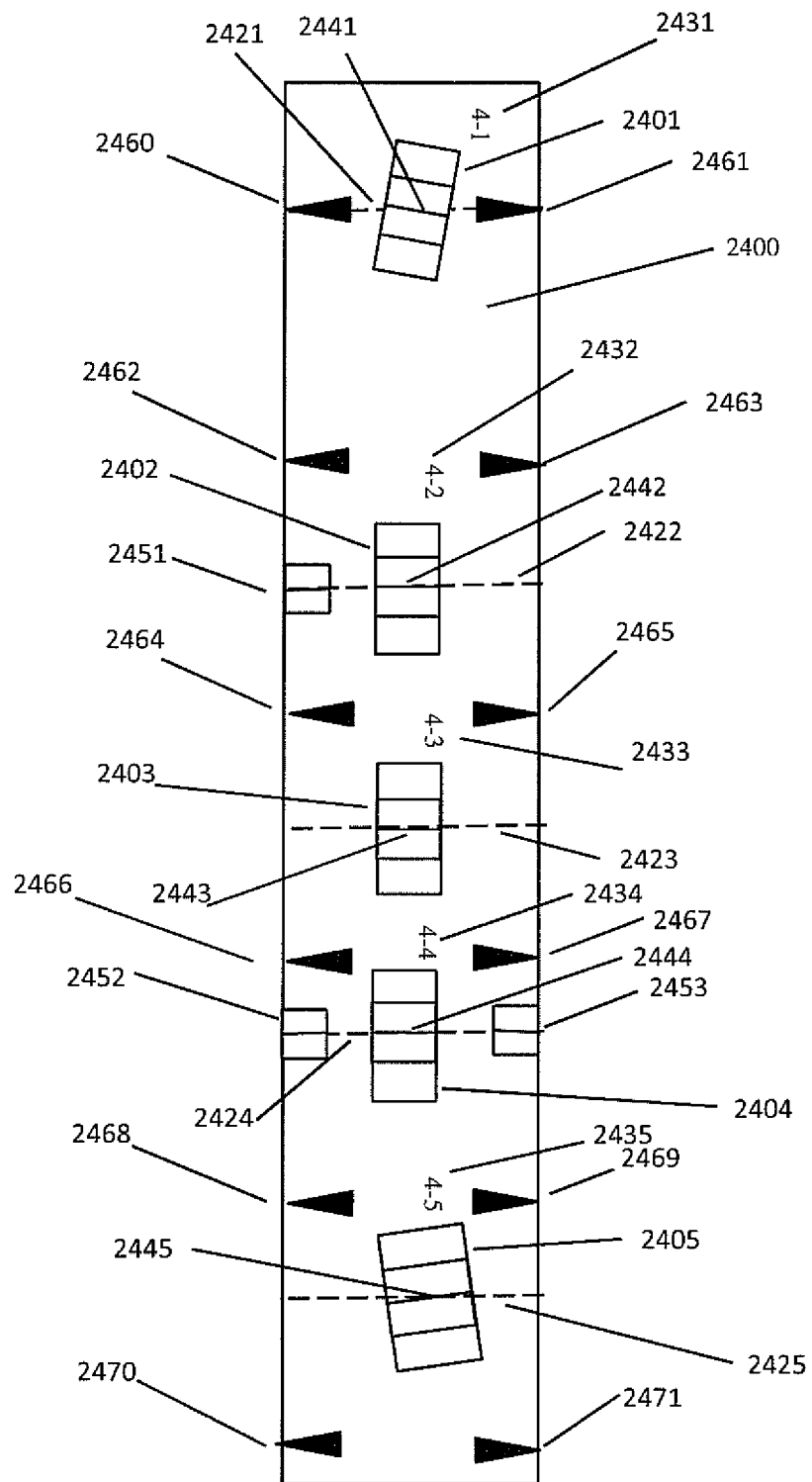


FIG 25

Target Number	Number Of Stripes	Color Pattern	Interval On Right	Rotation (in degrees)	Area
1	2	R/R	5	0	2x1
2	3	R/G/B	7	-20	2x1
3	3	G/B/G	12	0	4.5x1.5
4	3	R/G/B	4	-20	3x1
5	4	R/B/R/G	6	10	4x1
6	2	B/R	4	0	2x1
7	3	B/R/B	0	0	3x1

FIG 26

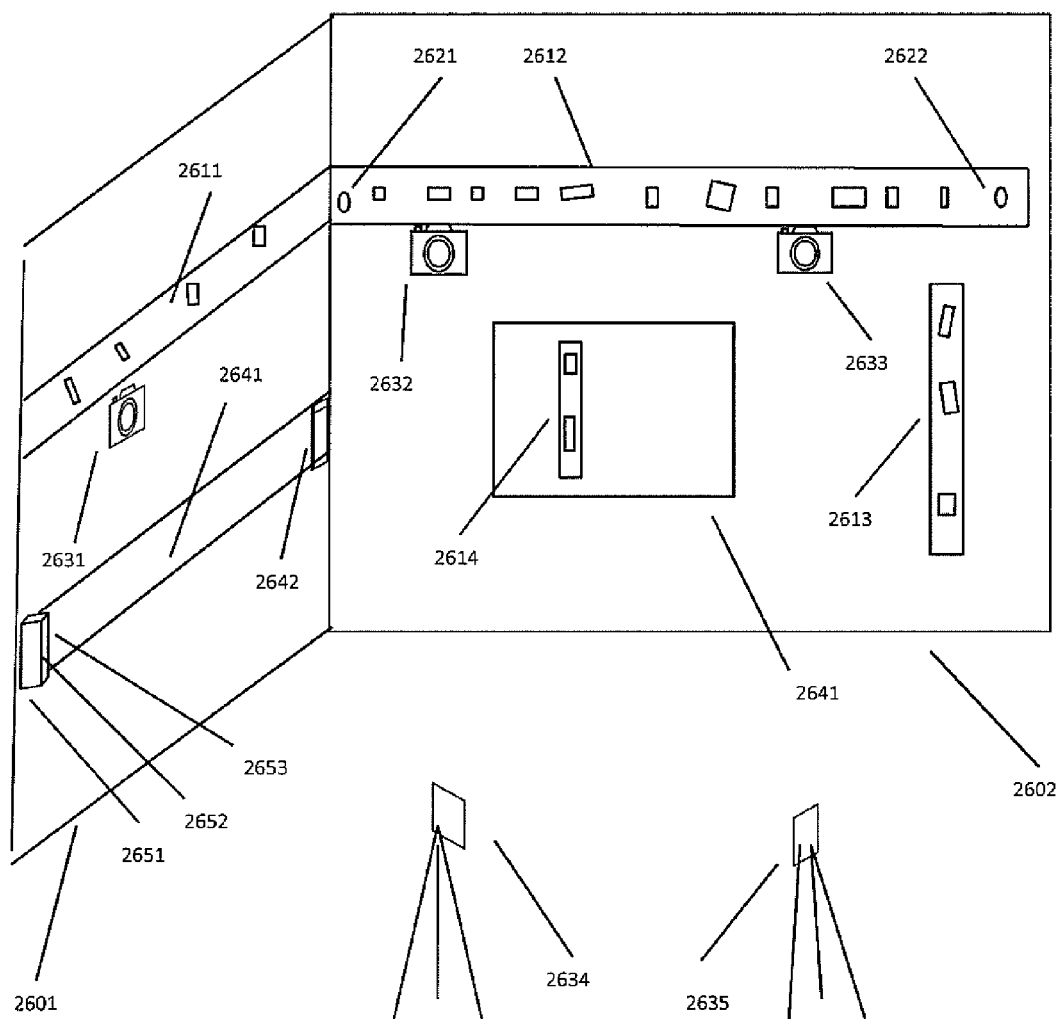


FIG 27A

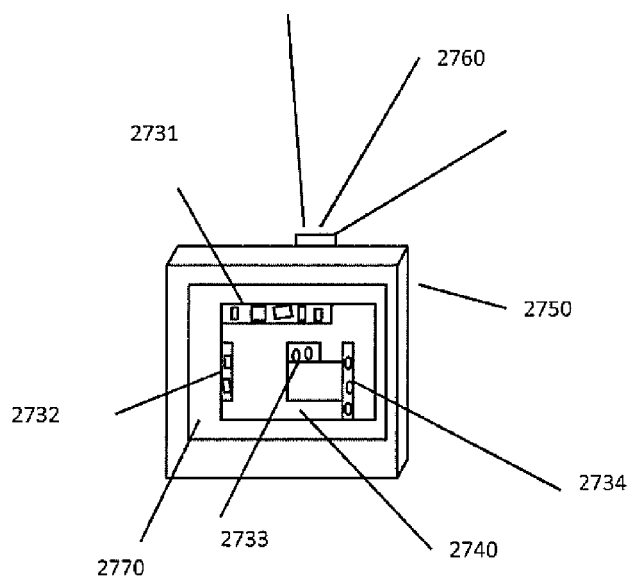
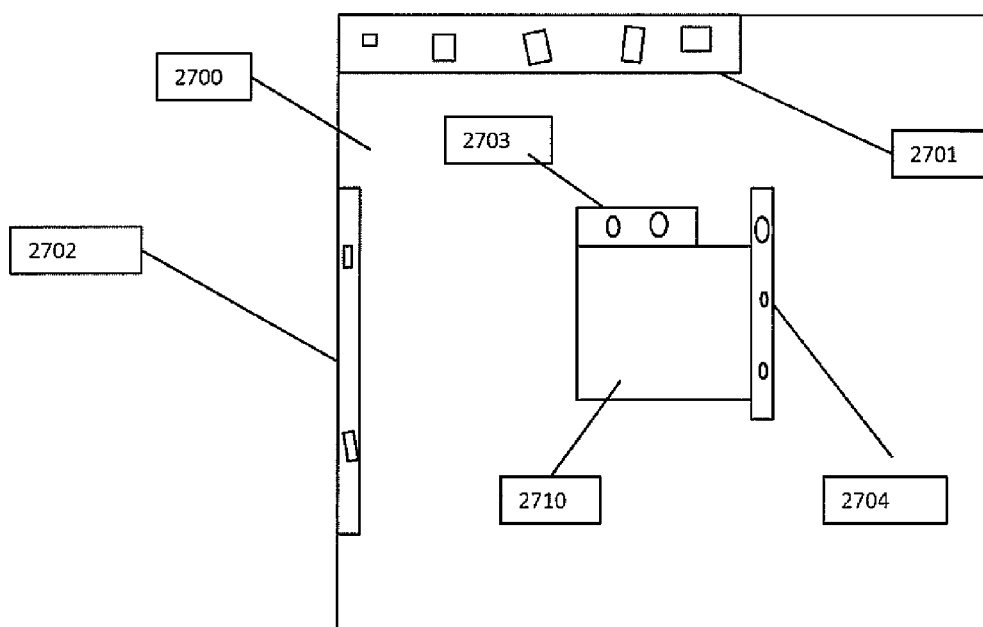


FIG 27B

FIG 28

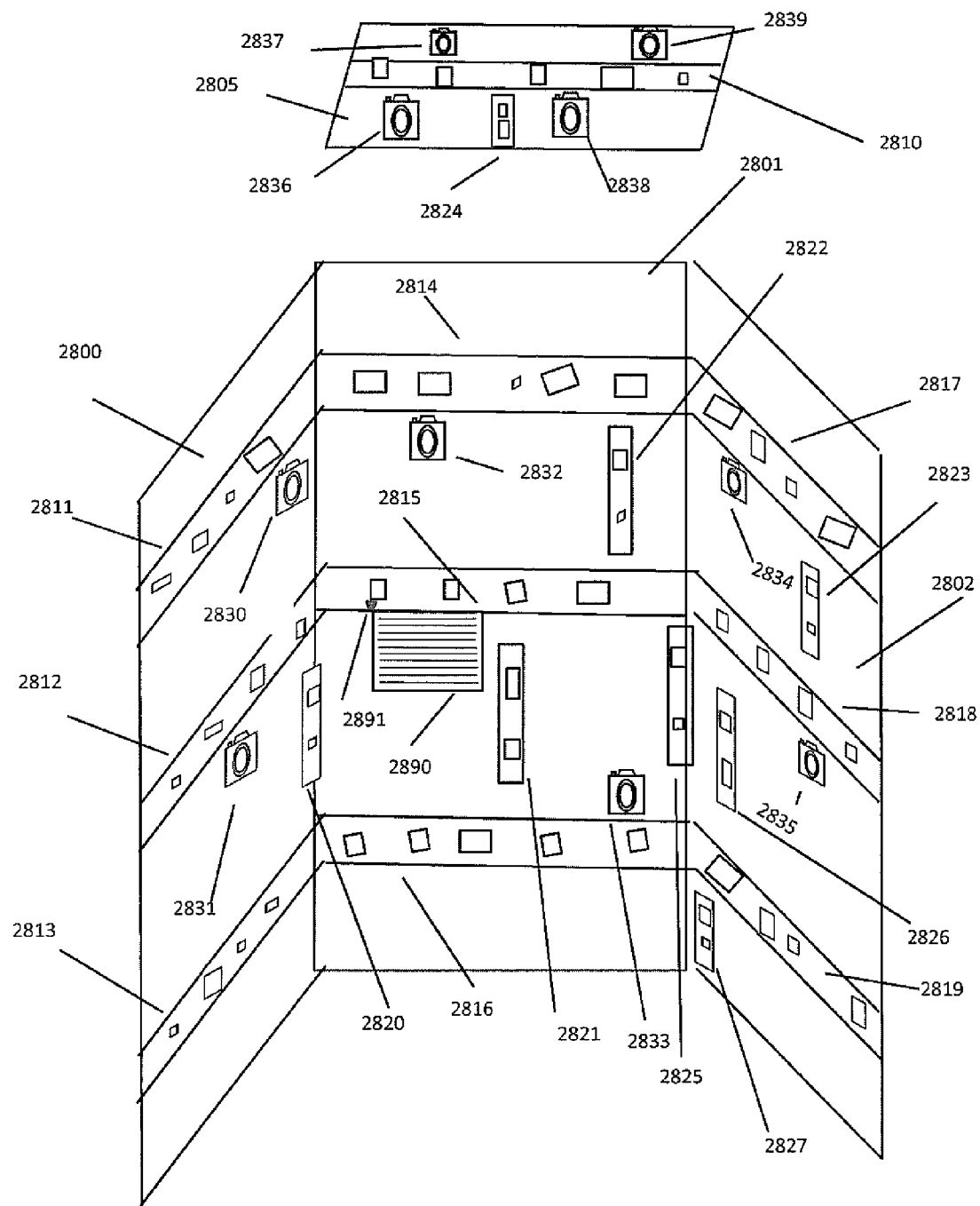


FIG 29

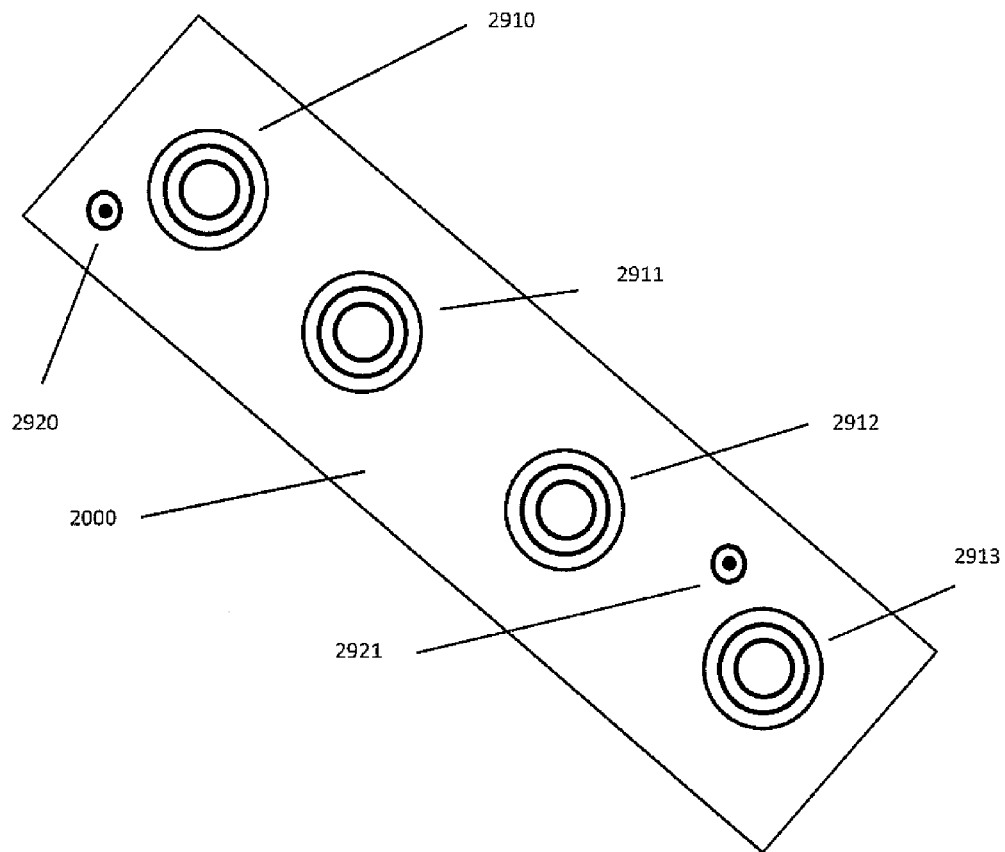


FIG 30A

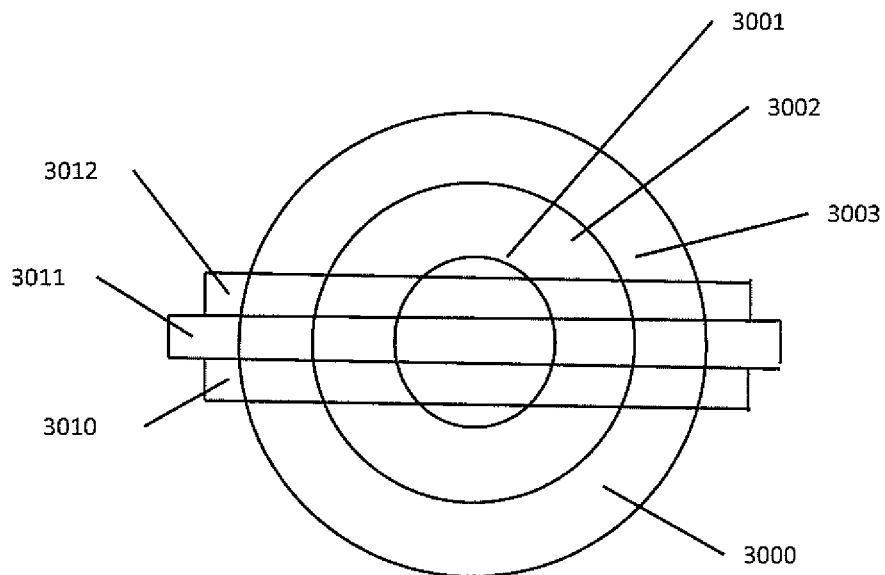


FIG 30B

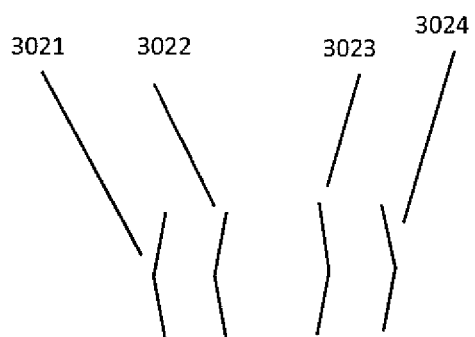
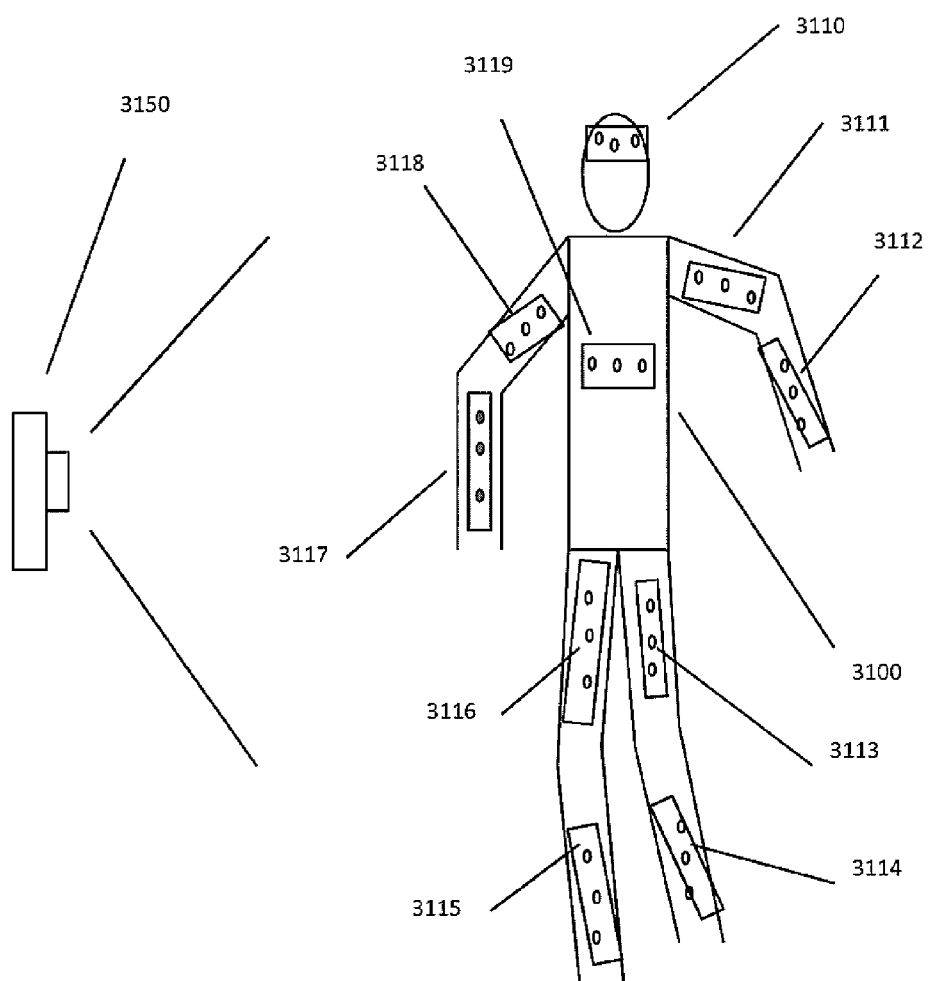


FIG 31



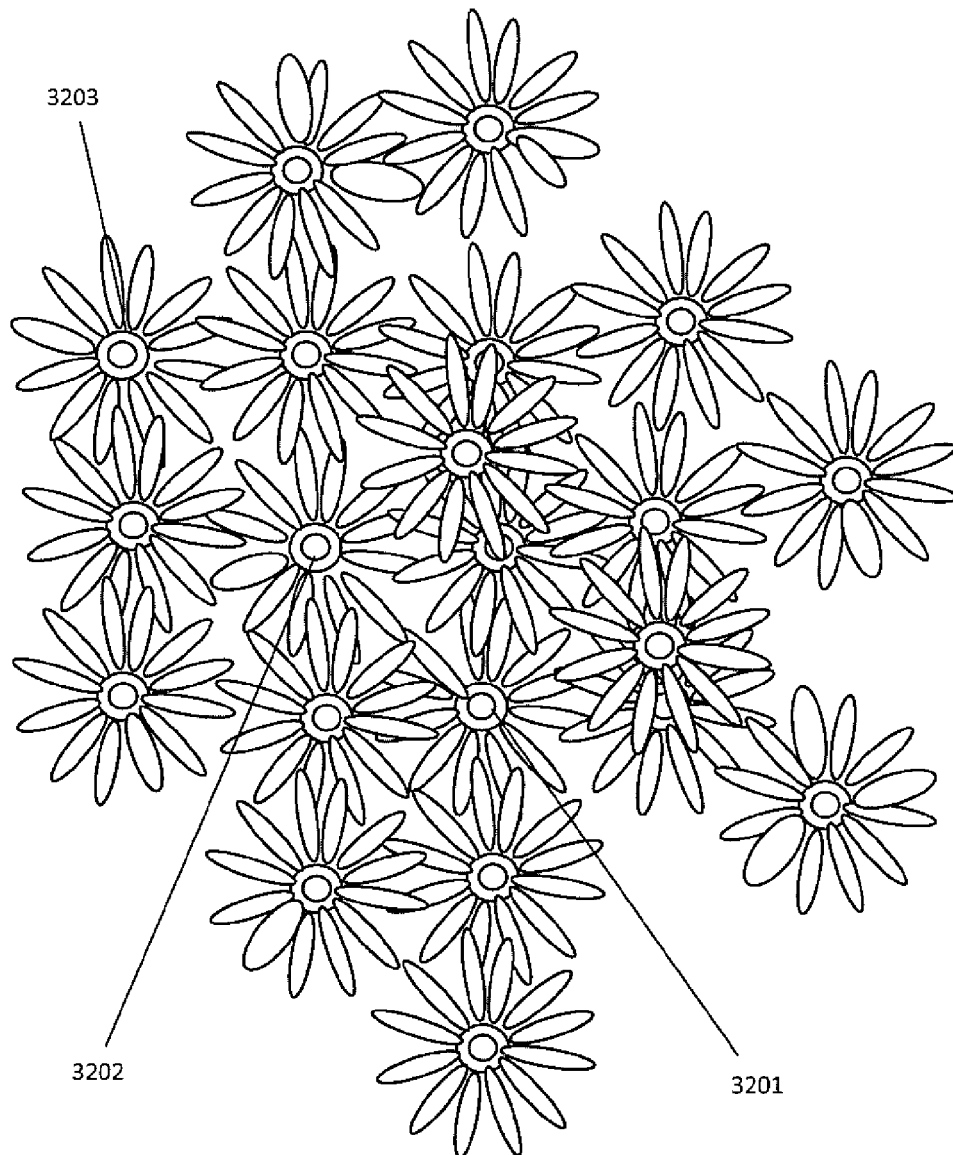


FIG 32

FIG 33

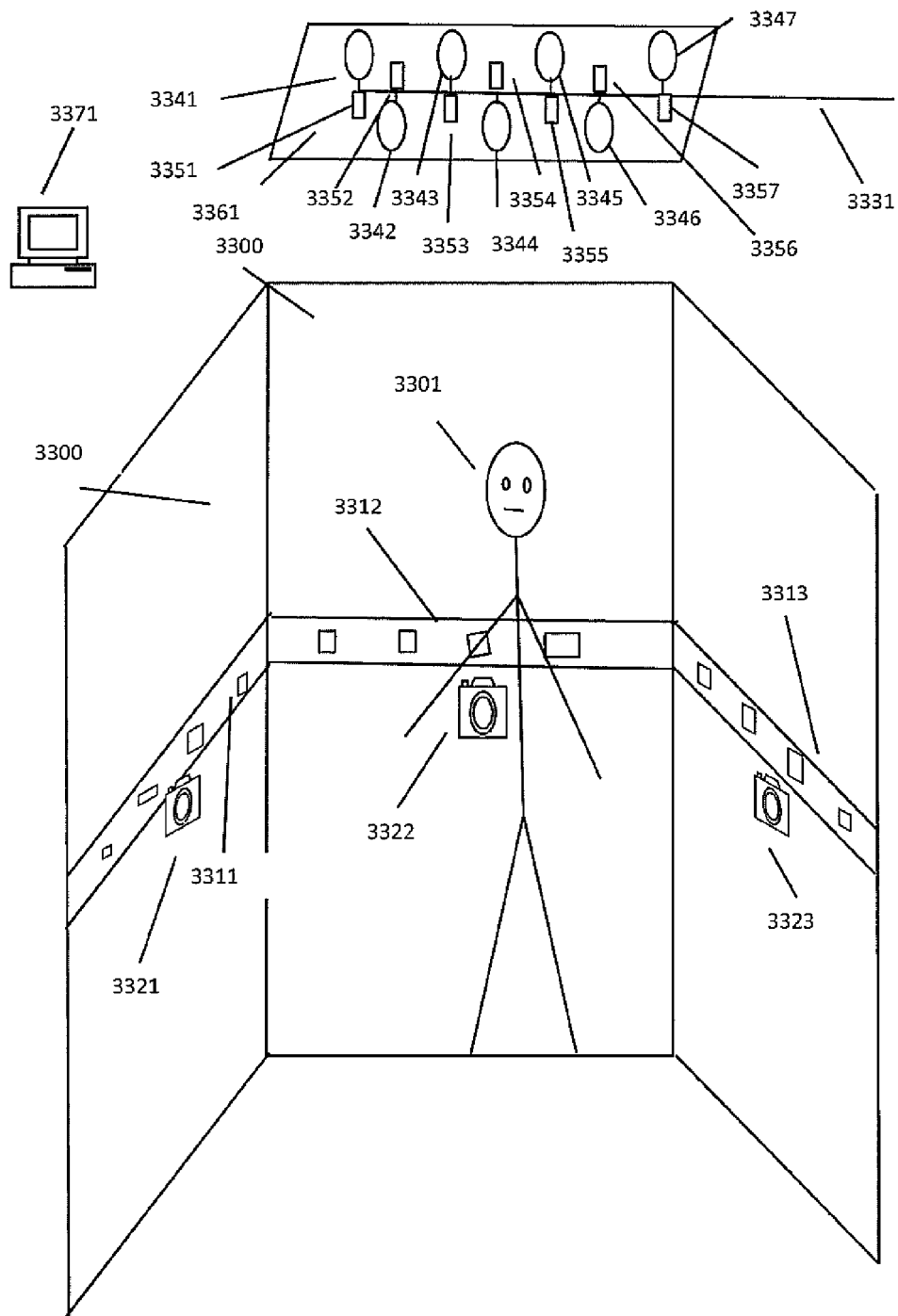
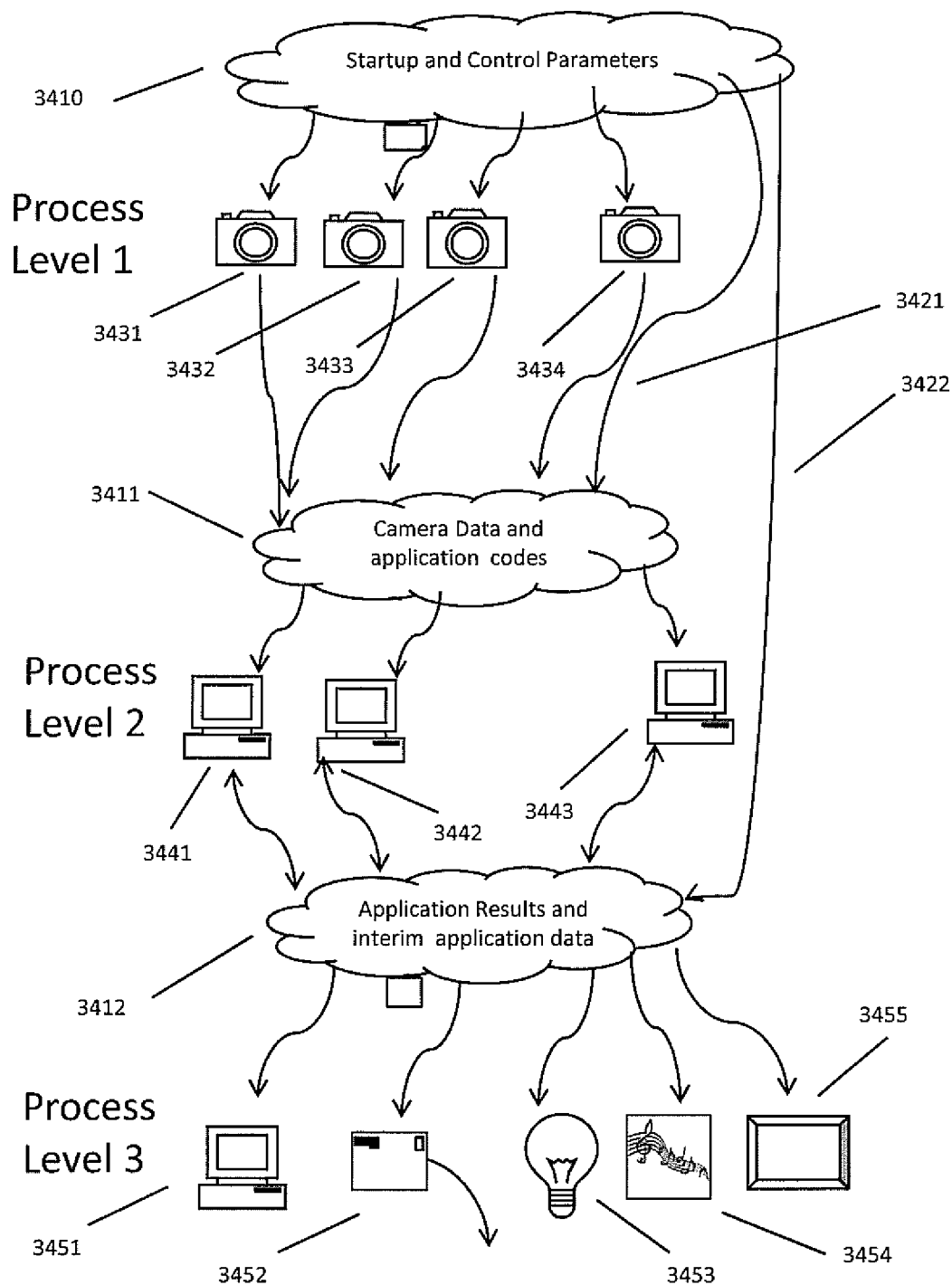


FIG 34



IMPLEMENTING RICH COLOR TRANSITION CURVE TRACKING FOR APPLICATIONS

CROSS REFERENCE TO CO-PENDING APPLICATIONS

This application is a continuation-in-part of co-pending U.S. patent application Ser. No. 13/301,216 filed on Nov. 21, 2011, which was a continuation-in-part of U.S. patent application Ser. No. 12/107,092 filed on Apr. 22, 2008, now U.S. Pat. No. 8,064,691 issued on Nov. 22, 2011, which claims priority benefit to the filing date of U.S. Provisional Patent Application Ser. No. 60/917,966, filed May 15, 2007, the contents of all of which are incorporated herein in their entirety.

BACKGROUND

The present disclosure relates, in general, to methods for finding targets and objects in an image.

As described in currently pending U.S. patent application Ser. No. 12/107,092 filed Apr. 22, 2008, in which the present inventors are named co-inventors, the region of interest (ROI) is the image currently being examined or possibly a subset of the image that is currently being examined. Within the ROI, the aforementioned method divides the pixels into Rich Color and non-Rich Colors. Rich Colors are further divided into Target Colors (these are limited to Rich Colors) and non-Target Colors. Target Colors are those Rich Colors that are of specific interest to the viewer at the time of examination. Rich, non-Rich, Target and non-Target Colors are reduced to binary representations, meaning that a type of color is represented as either a "1" or a "0". As described in the previous patent application, the use of configurable thresholds allows the user to simply control whether certain pixels are evaluated as target or non-target. Pixels which have been evaluated as being target pixels and that are adjacent to each other in the image are referred to as Rich Target patches. Pixels which are evaluated as being non-Target pixels can be used as a mask to restrict the area under examination reducing the amount of processing power and time required to process the ROI.

Typically, Blob analysis is used to find targets in a filtered image. Blob analysis is compute-intensive for complex images. The speed of the analysis is greatly reduced if the filtered image has image noise or a large amount of irrelevant objects. Roughly speaking, the time to process a set of blobs increase geometrically. This puts pressure on the user to better filter the image and to limit the size of the blobs to be tested such that there is an upper and lower limit to given characteristics. Further there is a need to restrict the search area to where it is most likely, i.e. near its last location. The targets also have limits as to out of plane angle that they can be tilted.

Traditionally developers of vision systems have sought to place limits on the targets, and the operational environments to improve the odds of search procedures. However, these limits greatly reduce the usefulness of the Blob method in everyday consumer applications. Ordinary users cannot be counted on to limit the size or distance of a target. They often move a target in and out of the cameras field of view ruining any chance of limiting the search region to a small fraction of the image. Consumer applications might have one target or ten or more. The number of targets could vary from frame to frame. Consumer applications are a nightmare. Our method does have a drawback also. This method can perform filtering in-line with Target detection. Consumer applications require inexpensive cameras and computing equipment that rapidly

and reliably tracks in normal lighting or poorly lit rooms with uneven lighting and in the hands of novice operators with little patience.

Machine vision, commonly called automated inspection, has been used in manufacturing processes to improve productivity and quality. On a typical production line, a sensor detects a part and signals a video camera positioned above or to the side of the inspection point to capture an image and send it to a machine vision processor. Using a combination of machine vision software and hardware the vision system analyzes the image and provides mathematical answers about the part. A traditional gray scale machine vision technology makes decision based on 0-256 shades of gray. A typical vision algorithm segments an image into pixels that fall within an intensity band bounded by a lower and upper threshold from the irrelevant pixels that have intensities outside of this intensity band. Alternatively they look at the rate of change of the image pixels. Once the relevant pixels have been identified, adjacent pixels are clumped together to form blobs and these are then characterized by geometric characteristics such as location, size, shape, etc. Inspecting colored parts or objects with gray-scale machine vision systems becomes usually unreliable in many cases and impossible in others. For this reason use of a color machine vision technology is needed to inspect parts or objects in ways that could not be done using traditional gray scale machine vision systems.

Thus far, color machine vision systems have been used for three primary vision applications:

Color Matching—verifying that a certain part's or object's color matches what the vision system is programmed to find:

Color sorting—sorting parts or objects based on color.

Color Inspection—inspecting colored parts or objects for defects or imperfections that gray scale image processing tools can't detect.

Defined as the perceptual result of visible light reflected from an object to human eyes, color represents an interpretive concept. Depending on how light is reflected, all humans see colors a bit differently. Human visual system use color to draw conclusions about surfaces, boundaries, location, and relative location to other objects, orientation, movement and changes in movement of objects in a scene. The human eye is usually capable of discerning both the color of objects under inspection and discerning the Transition Curves of said objects. Both Transition Curves and colors are used in building the "scene" that the brain uses to identify what is being viewed and to then make an interpretation of the meaning of what is seen.

Machine vision systems have typically reduced color information to one of 255 colors of gray scale in order to simplify processing. An undesirable byproduct of this simplification process is often the loss of important information which often reduces the utility of the inspection.

Color derives from the spectrum of light (distribution of light energy versus wavelength) interacting in the eye with the spectral sensitivities of light receptors. Typically, a wavelength spectrum from 380 nm to 740 nm (roughly) of light is detectable by human eye. This range is known as the visible light. The pure "spectral colors" from a continuous spectrum can be divided into distinct colors: violet (380-440 nm), blue (440-485 nm), cyan (485-500 nm), green (500-565 nm), yellow (565-590 nm), orange (590-625 nm), and red (625-740 nm). However, these ranges are not fixed, the division is a matter of culture, taste, and language. For instance, Newton added a seventh color, indigo, as wavelengths of 420-440 nm between blue and violet, but most people are not able to distinguish it. Of course, there are many color perceptions that by definition cannot be pure spectral colors. Some

examples of non-spectral colors are the "achromatic colors" (black, gray, and white) and colors such as pink, tan, and magenta.

An additive color system involves light "emitted" from a source or illuminant of some sort such as TV or computer monitor. The additive reproduction process usually uses red, green, and blue which are the "primary colors" to produce the other colors. Combining one of these primary colors with another in equal amounts produces the "secondary colors" cyan, magenta, and yellow. Combining all three primary lights (colors) in equal intensities produces white. Varying the luminosity of each light (color) eventually reveals the full gamut of those three lights (colors).

Results obtained when mixing additive colors are often counterintuitive for people accustomed to the more everyday subtractive color system of pigments, dyes, inks, and other substances which present color to the eye by "reflection" rather than emission. Anything that is not additive color is subtractive color.

Light arriving at an opaque surface is either "reflected", "scattered", or "absorbed" or some combination of these. Opaque objects that do not reflect specularly (that is, in a manner of a mirror) have their color determined by which wavelengths of light they scatter more and which they scatter less. The light that is not scattered is absorbed. If objects scatter all wavelengths, they appear white. If they absorb all wavelengths, they appear black. Objects that transmit light are either translucent (scattering the transmitted light) or transparent (not scattering the light).

The color of an object is a complex result of its surface properties, its transmission properties, and its emission properties, all of which factors contribute to the mix of wavelengths in the light leaving the surface of an object. The perceived color is then further conditioned by the nature of the ambient illumination, and by the color properties of other objects nearby, and finally, by the permanent and transient characteristics of the perceiving eye and brain.

Light, no matter how complex its composition of wavelengths, is reduced to three color-components by the eye. For each location in the visual field, the three types of color receptor cones in the retina yield three signals based on the extent to which each is stimulated. These values are sometimes called "tristimulus values".

To analyze and process images in color, machine vision systems typically use data from color spaces such as RGB, HSI (or HSL), HSV (or HSB), CIELAB (or CIEXYZ) CMYK, etc. Individual color within each of these spaces is sometimes referred to as a color component. The original color components can be scaled individually in either a linear or non-linear fashion before proceeding with this method emphasize given target characteristics or to take compensate for lighting or camera problems or even to implement the Rich Color Filter in a more efficient fashion.

In the RGB color space, each color appears in its primary spectral components of red, green, and blue. When combined with a three-dimensional coordinate system, the RGB color space defines quantitatively any color on the spectrum. RGB uses "additive" color mixing. X-axis specifies the amount of red color, Y-axis specifies the amount of green and the Z-axis Specifies the amount of blue. If RGB color model is implemented in 256 (0 to 253) discrete levels of each color component (8 bits) then the color space defines a gamut of $256 \times 256 \times 256$ or about 16.7 million colors.

The HSI color space, also known as HSL, is broken down into hue, saturation and intensity or lightness. Hue refers to pure color, saturation refers to the degree or color contrast, and intensity refers to color brightness.

HSV (hue, saturation, value), also known as SHB (hue, saturation, brightness), is quite similar to HSL "brightness" replacing "lightness". Artists often use HSV color space because it is more natural to think about a color in terms of hue and saturation.

CIE 1931 XYZ color space is the first attempt to produce a color space based on measurements of human color perception. It is the most complete color space used conventionally to describe all the colors visible to human eye. It was developed by the "International Commission on Illumination" (CIE). CIE 1976 LAB is based directly on CIE 1931 XYZ color space as an attempt to make the perceptibility of color differences linear. CIE is the most accurate color space but is too complex for everyday uses.

CMYK uses subtractive color mixing in used printing process. It is possible to achieve a large range of colors seen by humans by combining cyan, magenta, and yellow transparent dyes/inks on a white substrate. Often a fourth black is added to improve reproduction of some dark colors. CMYK stores ink values for cyan, magenta, yellow, and black. There are many CMYK color spaces for different sets of inks, substrates, and press characteristics.

Although dozens of defined color spaces exist, color machine vision applications primarily have used RGB and HSI or HSV color spaces.

Prior art systems use various techniques to measure and match colors such as a color sorting method for wires by comparing the output signal of a camera to the intensity ratio of known colors until a substantial match is found.

Another technique provides a color sorting system and method used for sorting fruits and vegetables. The sorting process is handled with a look up table. The pixel value of the input image is sent to the look up table and the output from the look up table is either series of 0's (accept) or 1's (reject).

Another method for automatically and quantitatively measuring color difference between a color distribution of an object and a reference color image uses "color distance" in a color system. A template representing the reference color image is stored in a memory of a machine vision system. The machine vision system generates a sample color image of the object and processes the template together with the sample color image to obtain a total color distance.

An apparatus is known for sorting fragments of titanium-based sponge on the basis color by comparing the color values of the image to a set of data values stored in a look up table for rejection or acceptance of each fragment.

Another system and method for locating regions in a target image matches a template image with respect to color and pattern information either by using a hill-climbing technique or fuzzy logic.

A different system and method of perceptual color identification can be used for the identification and tracking of objects, for example, in a surveillance video system. The described method includes a multilevel analysis for determining the perceptual color of an object based on observed colors. This multilevel analysis can include a pixel level, a frame level, and/or a sequence level. The determination makes use of color drift matrices and trained functions such as statistical probability functions. The color drift tables and function training are based on training data generated by observing objects of known perceptual color in a variety of circumstances.

It is clear from the prior art that traditional gray scale machine vision systems are being used successfully in a wide variety of inspection and process control applications for the electronic, automotive, food products, packaging, pharmaceutical, and recycling industries.

However, the use of color machine vision systems in these industries has only been applicable to well controlled immediate environments or surroundings. As machine vision is normally practiced, it is best to have the environment controlled in order to achieve predictable, high-quality results, for example:

Good lighting. The lighting should be consistent across the entire area being observed, without tints or shades (so that the true color of the objects under observation could be determined). When the case at hand requires that the ROI be observed multiple times the lighting should be maintained consistently over periods of time which could vary from milliseconds to months or years.

A simple, controlled, predictable background. Most systems require that the background of the objects under inspection be controlled and known in advance. For instance, a system designed to identify defective parts worked, in part, because the system knew it was always examining the parts against the consistent background of a black conveyor belt.

Control of items that are in the immediate area where the images are being gathered. Items positioned adjacent to the imaging area, or that move into or out of the imaging area during the time the images are being captured may affect the data that is captured by the cameras. This could be due to the uncontrolled items changing the amount or quality of light that is being reflected into or being absorbed adjacent to the imaging area or the color of light being reflected into the imaging area. For instance, if a person wearing a red shirt were to walk next to an imaging area during an image capture they could introduce additional red to the data gathered and potentially change how the data is interpreted.

Cameras (or other data gathering devices) that are of high quality, good color resolution, and that will produce repeatable data under similar conditions.

The cameras should be configured correctly (as to focus, shutter speed, aperture, sensitivity etc.).

The image should be captured with minimal motion blurring. This implies either that the items of interest should be held nearly motionless or that the camera(s) must be "fast" enough to capture a "frozen" image.

The orientation of the objects being examined should be known or controlled or compensated for along six axes—X, Y and Z along with roll, pitch and yaw. Lack of control of any of these, or lack of an ability to somehow determine these after the image is taken may render the image less useful. For instance, if a camera captures an image of a red box that is exactly four pixels high and four pixels wide it can make different deductions about the item depending on what it "knows" about that box. If it is known that the box is one mile away then one can reasonably estimate how tall and how wide the box is. Conversely, if it is known that the box is 10 feet tall and the image is four pixels by four pixels then the system can make a reasonably accurate estimate regarding how far away the box is.

Blob analysis is the most widely used tool to find and characterize objects and targets in an image. However, Blob analysis is very slow and requires large amounts of computer memory. The more complex the image or the more filter noise the slower the search. Finally, a successful blob analysis search requires that the size of the blob in the image does not vary widely.

These limitations make it difficult to build successful applications that use color cameras to control applications on common consumer devices such as cell phones or tablets where the user is a casual user, the equipment is inexpensive and the environment is often uncontrolled. Often the more

unsophisticated the user, the more demanding they are that technical products be fast, reliable, and simple to use with few if any restrictions.

It is also clear that prior art relied on matching color to a reference color image or template. A color machine and computer vision system that can make robust identification of color under varying lighting and changing image shift, scale, and rotation conditions is desirable. Machine vision systems use specialized and expensive hardware and software and therefore their use has been limited to industrial applications. With the advance of inexpensive color webcams, it is also desirable to find use for computer vision systems in cost sensitive consumer applications.

It would be desirable to provide a method to replace blob finding methods with a faster, more reliable method with fewer limitations for locating and identifying targets and objects in the field of view of a color camera in ordinary lighting environments, to thereby obviate for tracking purposes the prior art use of providing powered light sources in the target or in an illuminating source with specific directional or color characteristics.

It would also be desirable to provide a method that was so efficient that it could be implemented inside standard inexpensive cameras and that would transmit small packets of target data to a control center computer only when a target was identified.

It would also be desirable to provide an improved method of tracking objects that are tilted at extreme angles relative to the camera plane.

It would also be desirable to provide an improved method of tracking targets or objects through time or three dimensions.

It would also be desirable to track targets with patterns of Rich Colored patches.

It would also be desirable to provide an improved method for triggering interaction or applications between a user and computer by identifying, and locating targets or objects with Rich Colored patches in an image in the field of view of a camera or cameras.

It would also be desirable to provide an improved method for filtering such that everything in an image except for patches of different Rich Colors that were adjacent to each other was ignored.

It would be desirable to provide a machine or computer vision system that could be used in camera driven applications using ordinary smart phones, tablet computers, etc. that were fast inexpensive and reliable for unsophisticated users. It would also be desirable to provide a new method that is capable of incrementally processing a stream of data rather than requiring the processing of an entire ROI, that reduces or eliminates noise with little or no additional overhead, that discovers, assembles and stores geometric information that allows the determination of location, size, orientation, speed and acceleration of a target, and that is capable of this performance with a single target or multiple targets including tracking the position of multiple targets relative to each other within one ROI without severe degradation.

SUMMARY

In one aspect, a threshold color transition curve tracking apparatus is disclosed for use to identify a target in an image using a transition curve resulting from a threshold color transition points between two color types on a target placed on a surface and captured by an image capturing device.

The carrier can be an elongated strip.

The targets vary from each other in at least one of shape, color, patterns, area, angle of orientation on the carrier and the separation distance from adjacent targets on the carrier.

The targets can have one of a polygonal shape and a concentric circle shape.

The targets can be aligned along a centroid of the carrier.

The targets can include a first series of targets aligned along a centroid of the carrier; and satellite targets spaced from the centroid of the carrier.

The targets can be a multi-dimensional array of targets on the carrier.

The carrier can be a variable electronic display.

Location markers can be presented on the carrier at incremental dimensions along the carrier.

The colors on the targets can be rich colors, and the threshold color transition curves can be rich color transition curves.

In another aspect, a method is disclosed for identifying the location and orientation of an object in an image using a threshold color transition curve tracking apparatus to identify a target and an image using a transition curve formed of threshold color transition points between two color types on the target. The method includes providing a plurality of unique threshold color transition targets arranged in a unique sequence on a carrier, attaching the carrier to a surface, identifying the surface on which the unique arrangement of targets are attached by an image capturing device using the stored data defining the color transition curves of the unique sequence of targets to identify the surface on which the set of targets are protected.

In the method, the unique set of targets arranged in unique sequences are stored in a computer accessible memory. A computer processor compares a unique set of targets projected on one surface captured by the image capturing device with the stored data defining the unique set of targets to identify the target captured in the image.

The method uses a computer processor to identify color transitions exceeding a threshold between two color types on a target. The color types can be rich colors. The transition points in a target are reduced to a transition curve depicting the shape of the transition interface between two colors on a target.

This application teaches how a sequence of targets using Rich Color Transition Curves together with inexpensive equipment in natural lighting environments can be used to achieve results similar to those of the more traditional (and expensive) motion capture systems. We focus on motion capture because so much money and effort has been used to study how to do it and how to use it for other applications. However, the same methods can be used for applications that involve target tracking in one or more images.

In a previous patent U.S. Pat. No. 8,064,691, by the inventor, a method for finding Rich Colors in an image was disclosed. Among other things it defined a ratio space using the maximum component value for each pixel. This led to a distance equation in this ratio space that could be used to develop a test for the presence or absence of N factorial Rich Color types. If the number of color components $N=3$, then the number of Rich Color types is 6 (Red, Green, Blue, Magenta, Yellow, and Cyan). In some circumstances, black and/or white pixels are rare and, thus, worth defining as a Target color. The number of Target Color types is equal to or less than N factorial+2. Using a Rich Color distance equation, each pixel can be identified as having a Rich Color type and an associated Target Color ID value. If the pixel is not a Target Color, the Target Color ID is set to zero.

The image is processed along search-lines (rows, columns, etc.) initializing and updating a data set that is called a Rich Color Line-Set or just Line-Set defining a set of consecutive pixels all having the same Target Color ID and containing three pieces of data: the start pixel location of the pixel cluster, the end location, and the Target Color ID. Each time that the Target Color ID for the next pixel along the search-line is defined, it is tested to see if it has changed. If the Target Color ID is the same, then only the end location is updated to the new pixel location. If the Target Color ID has changed, a new Line-Set is defined with the location of the start and end both set to the new pixel location and the Target Color ID updated. Also, when the Target Color ID changes, begin testing for a Rich Color Transition by checking if there is the start location of the previous Line-Set is within a tolerance distance of another Line-Set further back in the current Search-line such that both have different but non-zero Target IDs. Any Line-Set further away than a distance tolerance, T_{LD} , is usually purged from memory to minimize memory requirements. If a Line-Set is found within the distance tolerance, T_{LD} , then a Transition Point Data-Set is created or updated. A Transition Point Data-Set consists of the following: a Transition Point ID number, a Transition Curve ID number, a Back-link Transition Curve number, a search-line number, a Transition Point location (X) on the search-line, a row/column indicator, the left Line-Set, and the right side Line-Set. The location X is defined by a statistical measure of the distance between the nearest start and end locations of the two Line-Sets and any non-Target pixel colors that fall between these locations. Often a simple average of the end of the left and end location of the two patches is the most common way to define x. However, sub-pixel accuracy can be obtained by using a weighted sum of the product of the pixel locations within the band extending from the end and start location mentioned above times their respective original color components all divided by the sum of these original color components. The new Transition Point can be linked up to points on the most recent search-line to find any Transition Point Data-Set located on the previous search-line that is within a distance tolerance, T_{PD} , and also with the same Target Colors on the same sides. If a Transition point is found that passes these tests, the new point's Transition Curve ID is set to that of the matching Transition Point-Set of the matching Transition Point-Set. If no matching Transition point is found then the Curve ID is incremented to one higher than the highest curve number.

The resulting curves together with their flanking Rich Color patches can be compared to target properties in a target database. The absolute and relative size, and orientation as well as shape and color can be used to identify, locate, and orient the targets. A sequence of images together with associated database properties can be used to determine physical properties of the objects to which the targets belong.

Rich Color filtering and the Rich Color Transition Curve detection can be combined in a compact implementation that operates on a stream of pixels occasionally sending a pixel with attending color filtering information to a second process that uses database information and previous Rich Color Transition Curve pixel data to link up Transition Curve pixels into a Rich Color Transition Curve. It is like a factory vision system that searches for rejects and then in the rare event that it finds such an object that fails to meet quality standards, a pusher pushes it into reject bin. However, in this case the rare event is finding a Rich Color Transition Curve point and instead of pushing it into a reject bin, it is sent to another process to string it together into Transition Curves and save

any color and location data with it. These Transition Curves have shape, color pairs, orientation, location, and size (both relative and absolute).

By creating a set of Rich Color Transition Curves we have reduced the original problem to one of comparing a database of target properties to a small set of curves with simple geometric and color properties. The Rich Color Transition Curves can be used to detect, identify, locate, orient, and characterize a set of objects or targets. One or more of these Rich Color Transition Curves can be identified as being part of a given target. A set of such Rich Color Transition Curves can be used to characterize a target and find a target's location, and orientation (in 2D or 3D space). A sequence of video images can be used to define physical properties and further filter the data by enhancing image quality (averaging target data) for targets that have matching characteristics but possibly a different location and orientation and by discarding data that cannot change properties as seen from frame to frame.

Once a target has been identified along with its geometric and other properties, these can be used with a database and a lookup table to manipulate computer interfaces, computer programs/macros, or devices. Both hardware and software implementations are disclosed. A smart camera with this method implemented in hardware could analyze each frame while the next frame is being acquired and then transmit (wired or wireless) a small data set of Rich Color Transition Curves only if a target is detected in the frame. This method requires only a few buffer memories and a very few multiples, adds, if tests and bit manipulation per pixel processed

The methods speed and flexibility make it possible to use a variety of targets at the same time. This method can be used to develop both 2D and 3D information. A sequence of video frames together with database information can be used to determine velocity, acceleration, and physics of a target object. Using targets with unique color and geometry, each target that is found can trigger a unique action based on their absolute or relative location and orientation together with their associated database properties.

To overcome these limitations related to capturing usable images for further processing, specific purpose-designed colored targets may be attached to items under investigation or known and controllable areas of an item may be specifically used as a target area (for instance a car assembly line may stop a car body under a camera at a specific point on the line with perfect lighting in order to verify that paint has been properly applied).

Purpose-designed colored targets provide unique functionality since their characteristics are known. The characteristics of a purpose designed Rich Color target may be stored in a database and made available to the image processing algorithm. The database about these targets can contain specific information such as the precise actual height and width of color patches on the target, the true colors of color blobs on the target, the relative location of Rich Color patches on the target, the relative location of multiple color patches on the target relative to all other color patches on the target, and information about Rich Color Transition Curves and boundaries on the color target including the length, position, absolute angles and relative angles of a Transition Curves to other Transition Curves and color patches on the target.

Targets, once located, can be used to trigger other actions. The mere presence or absence of a certain target, the orientation of that target, the movement pattern of that target, the speed of that target, or the acceleration or deceleration of the target may be used to trigger an action, program macro or some other external event.

Optical motion capture rooms often have uniform black or green walls/ceilings/floors and the actors have uniform colored suits with infrared reflectors on joints and fixed elements such as guns. This greatly cuts down the clutter of pixel data that must be tracked and analyzed by the computer.

The Rich Color Transition Curve method allows inexpensive applications to exist in natural environments with passive targets to achieve the same or better results than can be achieved in specialized motion capture rooms with infrared illuminators, retro reflectors, or high tech equipment. For many applications simple computing devices such as cell phones or tablets will do.

A 1D or 2D array of Rich Color Transition Curve targets can be attached to, painted on, or otherwise displayed on a surface with such that each target can be uniquely identified as well as located and oriented relative to other targets in the array. The most convenient implementation is to arrange a linear array (a sequence) of targets on a ribbon or tape.

BRIEF DESCRIPTION OF THE DRAWING

The various features, advantages and other uses of the present tracking for applications method and apparatus for implementing the Rich Color Transition Curve for tracking applications will become more apparent by referring to the following detailed description and drawing in which:

FIG. 1 illustrates a computer vision system which performs color identification;

FIG. 2 illustrates a computer vision system that performs color tracking;

FIG. 3 illustrates the components of a typical computer system used to implement the software version of this method;

FIGS. 4 and 5 are pictorial representations of Rich Color Transition Curve tracking according to using a Tablet computer or Smart Phone;

FIG. 6 is a pictorial representation showing the use of cameras implementing the Rich Color Transition Curve method in vision applications;

FIG. 7 is a pictorial representation showing how a module containing circuitry utilizing the Rich Color Transition Curve method can be used to upgrade standard video cameras;

FIG. 8 is a pictorial representation showing image data coming from a data storage device to be analyzed by the Rich Color Transition Curve tracking;

FIG. 9-19 are pictorial representations illustrating the progressive results of the processing steps of the Rich Color Transition Curve method from an original image with multiple objects and targets with color components to a single sought after target;

FIGS. 20 and 21 are flowcharts depicting the sequence of steps in the present method as performed one pixel at a time on a data string; and

FIGS. 22 and 23 are pictorial representations showing examples of data sets used to find transition points, build transition curves, and find targets.

FIG. 24 is a pictorial representation of an example of tape implementation of Rich Color Target Sequences;

FIG. 25 is a sample table showing the contents of a target sequence database.

FIG. 26 is a pictorial representation of an example of an optical motion capture room using rich color target sequences;

FIG. 27 is a pictorial representation of a method showing how to determine the relative location of an object using an image on a tablet screen;

11

FIG. 28 is a pictorial representation of a room that is setup for automated calibration of the room using Rich Color Target Sequences;

FIG. 29 is a pictorial representation of an example of a Bulls Eye Target Sequence.

FIGS. 30A and 30B are pictorial representations illustrating how a horizontal scan line would view a Bulls Eye Target using Rich Color Transition Curves;

FIG. 31 is a pictorial representation illustrating an actor wearing Rich Color Target Sequences for motion capture;

FIG. 32 is a pictorial representation illustrating how a Rich Color Target can be hidden from human vision using an artistic display and yet be the only thing that a Rich Color Transition Curve method can see;

FIG. 33 is a pictorial representation of an example of an application that would best be performed by hardware implementation of Rich color Transition Curve methods Example of Rich Color Transition Curve hardware in order to preserve the modesty of a user in a wireless age; and

FIG. 34 is a block diagram illustrating how a set of repeatable hardware and software components can be used in a three-step process tied together with cloud communication to solve a variety of machine vision tasks for rigid bodies using Rich Color Transition Curve methods with Rich Color Target Sequences.

DETAILED DESCRIPTION

FIG. 1 illustrates one aspect of a computer vision system 10 that performs color identification. The computer vision system 10 may include a computer system 11, a color camera 12, such as a webcam, and a field of view 13. The computer system 11 may include one or more processors, a memory medium, monitor, and input devices, such as a keyboard and mouse and any other components necessary for a computer system. The computer system 11 also includes one or more software programs operable to perform color identification function. The software programs may be stored in a memory medium, such as a DRAM, SRAM, EDO RAM, etc., or a magnetic medium such as a hard drive, DVD, CD, or floppy disk. The computer system 11 is broadly defined to encompass any device, having a processor which executes instructions from a memory medium, such as a personal computer, workstation, mainframe computer, network appliance, internet appliance, personal digital assistant (PDA), cell phone, iPod, etc.

The color camera 12 can be an inexpensive webcam. The color camera 12 may comprise an image sensor such as a "Charged Coupled Device" (CCD) or "Complementary Metal Oxide Semiconductor" (CMOS). The color camera 12 may be connected to the computer system 11 USB port either through a wire or wirelessly. The cameras and the computer do not have to be collocated; they might even be 2000 miles apart. The color camera 12 may be attached to a flexible stand or clipped on a monitor to point at a particular field of view 13. The output of the color camera 12 is usually the values in 256 discrete levels of each of three color-components, red, green and blue (R, G, B), for each pixel of a target image in the field of view 13. The pixel-by-pixel color information of the target image is fed to the computer system 11 for each frame and this information is repeated on a continuous basis depending on the refresh rate of the color camera 12. The way the color information is processed by the software program of the computer system 11 is explained in details below.

The color identifying method can identify six (three factorial) colors; red, green, blue, yellow, cyan, or magenta with the use of three-component color camera 12 as well as black

12

and white for a total of eight colors. With the advance of the four-component color cameras, 24 (four factorial) colors or a total of 26 colors including black and white can be identified. The present method identifies the colors of interest on a target image accurately under varying light and image conditions.

As a first step, the method receives the output information of the camera expressed in (R, G, B) values of color components of each pixel. The largest color component is then identified and all three color-components (R, G, B) are divided by this value. It is important to note that the largest color component may be different from pixel to pixel and is not an overall or fixed maximum. In this way, the present method creates a new color space called "Ratio Space". The components of the ratio space (r, g, b) are such that the largest component is always and the other two components may be 0 or 1.0 or a value between 0 and 1.0.

From this point on, the method processes the color information from each pixel in ratio space values (r, g, b). Next, the ratio space values (r, g, b) are put to a "Threshold Test". If the values pass the threshold test then the information is identified as a "rich" shade of the color of interest. The present method departs from the prior art in that the prior art tries to identify every shade of a color on the target image by matching that color to an elaborate library of reference color images or templates. The improved method effectively and accurately identify "rich" shades of a color of a target image from the "pale" shades of a color under varying light and image conditions. Once the relevant pixels are identified as "rich" shades, the adjacent pixels are clumped together to form blobs and these blobs are then filtered by geometric characteristics such as shape, size, location, orientation, etc.

The method then keeps track of the information of a target image from one frame to the next. Any changes in the target image from one frame to the next or succession of frames can be used as an interaction between the user and computer. This interaction can be in the form of performing certain tasks or initiating applications or feedback, thus making the camera a convenient interface for the user. Thus, the first step in tracking is filtering out of the clutter of the target image all but a specific rich color. Next, this simple image is filtered to find blobs of this color with specific shape and size. This step is repeated for other specific rich colors. And finally, a target or set of targets of that are geometrically related to each other can simply be identified and used to trigger a computer action.

The threshold test is carried out in a "Distance" equation defined below. The distance equation converts color information from each pixel, in ratio space values (r, g, b), to "achromatic" color information (black, gray, or white) between 0 and 255 or more preferably to "binary" information black or white (0 or 255). The method creates a "Filter" by combining the threshold test into the distance equation and accomplishes to reduce the color information of a target image into a binary output, black or white. Black represents the color information that passed the threshold test as a "rich" shade of a color of interest or "target" and white represents the color information that failed the threshold test as a "fade" shade of a color or "unidentified" color. Thus, with a three-component color camera, the method can separate a target image into 6 regions of distinct colors.

The distance equation employs a "Scale Parameter" (S). The scale parameter is usually a very large number and set to a "negative" value for the primary component(s) of the color of interest so that it operates in the opposite direction to the "Threshold Value" (T). The distance equation also employs a

13

function called POS (q) and $\text{POS}(q)=0$ if $q \leq 0$ else $\text{POS}(q)=q$. The distance equation is defined as follows in terms of ratio space color component values (r, g, b):

$$\text{Dist} = \text{POS}(Sr * (r - Tr)) + \text{POS}(Sg * (g - Tg)) + \text{POS}(Sb * (b - Tb))$$

Exemplary threshold values and scale parameters for 6 colors of interest are as follows:

RED: $Tr=1.0$, $Tg=0.8$, $Tb=0.8$ $Sr=-1000$, $Sg=1000$, $Sb=1000$

GREEN: $Tr=0.8$, $Tg=1.0$, $Tb=0.8$ $Sr=1000$, $Sg=-1000$, $Sb=1000$

BLUE: $Tr=0.8$, $Tg=0.8$, $Tb=1.0$ $Sr=1000$, $Sg=1000$, $Sb=-1000$

YELLOW: $Tr=0.95$, $Tg=0.95$ $Tb=0.8$ $Sr=-1000$, $Sg=-1000$, $Sb=1000$

MAGENTA: $Tr=0.95$, $Tg=0.8$, $Tb=0.95$ $Sr=-1000$, $Sg=1000$, $Sb=-1000$

CYAN: $Tr=0.8$, $Tg=0.95$, $Tb=0.95$ $Sr=1000$, $Sg=-1000$, $Sb=-1000$

The method can also determine the achromatic colors such as black and white when all three color components in ratio space (r, g, b) are 1.0 or nearly 1.0, if so by looking at the original (R, G, B) values being (large) above a white threshold or (small) below a black threshold.

For a given pixel of color information, if the output of the distance equation is 0 then that color passes the threshold test, if the output of the distance equation is anything but 0 then that color fails the threshold test.

The following example demonstrates how distance equation filters the color information from the camera output to binary color information:

Example 1

Consider two pixels with the following components: Pixel 1: (R, G, B)=210, 50, 40 and Pixel 2: (R, G, B)=210, 190, 80 In ratio space values: Pixel 1: (r, g, b)=1.0, 0.238, 0.190 and Pixel 2: (r, g, b)=1.0, 0.904, 0.381 then the distance equation for the Pixel 1 and Pixel 2 become:

$$\text{Dist1} = \text{POS}(-1000 * (1.0 - 1.0)) + \text{POS}(1000 * (0.238 - 0.8)) + \text{POS}(1000 * (0.190 - 0.8)) = 0 + 0 + 0 = 0$$

$$\text{Dist2} = \text{POS}(-1000 * (1.0 - 1.0)) + \text{POS}(1000 * (0.904 - 0.8)) + \text{POS}(1000 * (0.381 - 0.8)) = 0 + 10.4 + 0 = 10.4$$

The result of distance equation is "0" i.e. the Pixel 1 passes the threshold test and is identified as a rich shade of red and the output of the filter is set to black. On the other hand, Pixel 2 does not pass the threshold test and is categorized as a fade or pale shade or unidentified color, therefore, the output of the filter is set to white (i.e. 255).

There are several ways for defining a filter and setting threshold values. For example, a pixel representing a green color might register the following values in the ratio space: (r, g, b)=0.45, 1.0, 0.55. A filter can be constructed such that anything with $Tr \geq (1.45/2)$ or $Tg \leq 1.0$ or $Tb \geq (1.55/2)$ is rejected by the filter. This threshold is called the "half-distance-value" to the primary color component (1.0).

The method can be enhanced to handle cameras that are not calibrated correctly for the ambient lighting. This requires a preprocessing phase that consists of the following steps: First, identifying the component bias of each color component (R,G,B). This can be done by red, green, blue targets or a set of known black blobs and identify the lowest component values of each of these colors. Subtract each of these three values from their corresponding component in each pixel of the entire image. Second, multiply each R,G,B value of every

14

pixel in the image by a single scale factor so that the entire image brightness is enhanced to compensate for the brightness that was subtracted. For the ratio signature space, this step is unnecessary since the ratio cancels out any factor that is common in both the numerator and the denominator.

To provide successful commercial applications in color identification, the method should be very robust in every lighting condition. A field of view might be under direct sunlight, in a shadowy room, or under incandescent lights during evening, etc. The strength of the method in identifying color particularly in challenging lighting environments comes from the "Ratio Space". The ratio space has an impact on finding targets and colored objects in a typical environment for commercial and consumer applications. The following example illustrates this point:

Example 2

The camera output might register (R, G, B)=0.6, 0.8, 92.8 and (r, g, b)=0.006, 0.008, 1.0 for a blue spot over a sunny part of the field of view or (R, G, B)=3.2, 14.3, 63.5 and (r, g, b)=0.05, 0.225, 1.0 over a shadowy region of the field of view. The camera output for a red spot might register (R, G, B)=99.6, 0.4, 0.4 and (r, g, b)=1.0, 0.004, 0.004 over a sunny part of the field of view or (R, G, B)=64.7, 17.8, 4.6 and (r, g, b)=1.0, 0.275, 0.07 over a shadowy region of the field of view. While the original (R, G, B) values might fluctuate significantly from sunny regions to shadowy spots of the field of view, the ratio space values make it easy to identify the color of interest.

Another advantage of the present method in identifying color is the ability to optimize the "camera parameters" for varying lighting conditions. Camera parameters such as: gain, brightness, contrast, saturation, sharpness, white balance, backlight compensation, etc. can be optimized for a given field of view and the accompanying lightning conditions. The method accomplishes this optimization by going through a calibration process for a known field of view as a preprocessing step. Once the camera parameters are optimized for a given field of view, the method is ready to launch.

The field of view 13 for the present method can be anything that the camera 12 is pointing at. The camera 12 can be pointing at a desktop such as in FIG. 1, and in this case, the field of view 13 can be a plain sheet of paper, a book, an object, etc. The camera 12 can be pointing at a person or people in front of the computer, or a scene with items or objects in it. The field of view 13 can be a screen or whiteboard that the camera 12 is pointing at. Further, the target image that is processed by this method can be the entire field of view or part of the field of view such as an "area of interest". For example, not every item or object in the field of view might be changing from one frame to the next. In this case, the target image might focus on the section of the field of view that might be an area of interest.

It should be by now obvious to one skilled in the art that the present method can be used in a variety of consumer and commercial applications. One aspect of creating consumer friendly applications using the method is the ability to identify color effectively under varying lighting conditions in the field of view of a camera. The monitoring and tracking changes in the field of view of a camera lead to potential uses not only in traditional machine vision applications but also open up consumer applications with the use of inexpensive webcams.

FIG. 2 illustrates a computer vision system that performs color tracking according to one aspect of the present method. An application of the present method is given in FIG. 2 as an

15

example for tracking an object or target **20**, such as a pen, in the field of view by identifying relative locations of colored sections of this object. Tracking of simple objects such as a pen or finger in a field of view can be used as an alternate input device for computer aided design and drawing CAD) applications.

Another aspect of the present method is based on locating a set of points that transition from a patch of one type of Rich Color to an adjacent patch of a different type. Whenever possible these points are strung together to form a chain which is referred to as a Rich Color Transition Curve. This method can be used to detect and locate targets in an image, to distinguish one target from another, to define target properties, or filter data. The image is typically, but not always, in the field of view of a camera or cameras. The target properties can be compared to a database and then used as an interface to a computer for machine and computer applications.

This aspect also relates to a triggering mechanism or mechanisms by identifying Rich Color Transition Curves, combinations of Rich Color Transition Curves and Rich Color boundaries along with their colors, orientations, positions, and motion (both speed and acceleration) combined with a computer or machine interface which allows the invention to be used for consumer or industrial applications

DEFINITIONS

Color component—the color components C_n where $1 \leq n \leq N$ that make up an image. In the most common case $N=3$ and C_1 =Red, C_2 =Green, and C_3 =Blue.

Ratio Color component—the Ratio Color components c_n for a pixel where $1 \leq n \leq N$ are obtained from the input colors C_n as defined by $c_n = C_n / C_H$ where C_H is the largest component of this pixel.

Type or ID—most of the terms defined in this section have an associated identifying number or ID.

Rich Color—A color made up of 1 or more color components that are much larger than the remaining color components as typically defined by the Rich Color Filter equation for a filter type “m” color distance equation. For a pixel located I, j the color distance equation is defined as follows in terms of ratio space color component values (c_1, c_2, \dots, c_N): $A_m = \text{POS}(B_{1,m} * (f(c_1) - T_{1,m})) + \text{POS}(B_{2,m} * (f(c_2) - T_{2,m})) \dots + \text{POS}(B_{N,m} * (f(c_N) - T_{N,m}))$

Where, for color component n and color filter type m , $T_{n,m}$ is a color threshold, $B_{n,m}$ is the Bias, and the function $\text{POS}(A)$ such that $\text{POS}(A)=0$ if A is less than or equal to 0, ELSE $\text{POS}(A)=A$. The function $f(c_n)$ can simply be any function that emphasizes Rich Color for example $f(c_n)=[E_1 * c_n]+E_2$ or it might be $f(c_n)=[E_1 * c_n * c_n]+[E_2 * c_n]+E_3$ for input parameters E_1, E_2, E_3 . Typical Rich Color Filter Parameters ($T_{1,m}, T_{2,m}, T_{3,m}, B_{1,m}, B_{2,m}, B_{3,m}$) for RGB images are as follows:

Filter type $m=1$, RED: (1.0, 0.8, 0.8, -1000, 1000, 1000)
Filter type $m=2$, GREEN: (0.8, 1.0, 0.8, 1000, -1000, 1000)
Filter type $m=3$, BLUE: (0.8, 0.8, 1.0, 1000, 1000, -1000)
Filter type $m=4$, YELLOW: (1.0, 1.0, 0.8, -1000, -1000, 1000)
Filter type $m=5$, MAGENTA: (1.0, 0.8, 1.0, -1000, 1000, -1000)
Filter type $m=6$, CYAN: (0.8, 1.0, 1.0, 1000, -1000, -1000).

Enhanced Ratio Color types—The total number of color categories, M , that include both Rich Colors, N factorial ($N!$), that can be developed from a Rich Color Filter plus Black, and White. Typically $N=3$ and $M=8$.

Target Color—a subset of L Rich Colors and possibly Black and/or White that are used to find a target or object in an image such that $L \leq M$.

16

Rich Color Boundary—The perimeter of a Rich Color patch that separates it from non-Rich colored areas.

Transition Curve—The curve or chain of points or polyline that represents the separation of two Rich Color patches and include Color Pairing data. For simplicity even a single point is referred to as a curve or chain of length one.

Color Pairing—The property of a Transition Curve which identifies a left side color and a right side color as a curve is oriented facing the end point.

Search-line—A row or column or angled line of pixels across the image. A complete set of consecutive search-lines defines the complete image.

Region of Interest (ROI)—rectangular subsection of the image to be studied

Transition Point—The location on a search-line that best defines the transition from a patch of one Rich Color type to a nearby patch of a different Rich Color type.

A block diagram shown in FIG. 3 for a typical computing system, which may include the computer system **11** shown in FIGS. 1 and 2, includes a central processing unit (CPU) or a processor **330**. Input/output (IO) circuitry **332** is coupled to the processor **330** and provides an interface to and from the processor **330** from inputs and outputs, such as a display/input tablet **334**, or a camera **336**. Internally, the processor **330** is coupled to a clock **333** and memory **331**, such as ROM memory and/or RAM memory.

The method is embodied in software code on a computer medium which may be portable or a computer medium attachable to the processor **330** for execution by the processor **330** or stored in one or both of the memories **331**. The memories **331** maybe external from the processor **330** integral with the processor **330**, etc.

The display/input table **334** may be a separate display and a separate tablet or input device, such as a keyboard, mouse, etc.

The display/input tablet **334**, whether embodied in a single integrated unit or in separate units, will have appropriate audio output and possibly a microphone input.

It should also be noted that the camera **336** maybe integrated as part of the display/input tablet **334** or as a separate element as shown in FIG. 3. Further, all of the computing device circuitry, such as a processor **330**, memory **331**, input/output circuitry **332** maybe incorporated internally as part of the camera **336**.

FIG. 4 illustrates how Rich Color Transition Curve tracking can be used to develop a Beverage Can Game using a standard tablet computer or smart phone **104**. The game, for example, increases or decreases the volume of audio channels for voice, drums, guitar, etc. by tying relative positions and orientations of cans held by different players. Then the tempo or other characteristics can be tied to the absolute positions of the cans. The can 3D positions and orientations (roll, pitch, yaw) are determined by tracking the length, orientation, and other geometric and motion properties from frame to frame of the camera using the methods described in this application. FIG. 4 shows two cans **400, 401** each with a unique target **405, 408** that is viewed by a camera **402** that is in a device with a computer (in this case a tablet computer **404** with display **403**). The target Rich Color patches are separated by Transition Curves.

FIG. 5 show the same set of objects one frame later where the cans **400, 402** have moved. The change in overall size of a target in the image can be used to determine the target's distance from the camera. Pitch and yaw can be determined by determining a preferential shrinkage of the target along a polar angle from the center of a target.

17

FIG. 6 illustrates how a camera 650 that has the Rich Color Transition Curve Method built in can be used in vision applications. Wireless communication 651 only has to ship a small set of Transition Curve data for each frame. The central processor or computer 604 can use the transition curves as target signatures and, when targets of interest are detected, a full analysis of the target location and orientation can be conducted.

FIG. 7 illustrates how a small module 750 housing simple circuitry embodying the Rich Color Transition Curve Method can be used to upgrade a standard video camera 751. The module 752 is wirelessly or hardwired to the computer 404. FIG. 8 illustrates how the Rich Color Transition Curve Method can be used on an image previously stored in a storage device 850 coupled by the cables 851.

One version of the beverage can game uses the cans to control the sound tracks that make up a song. For instance, the distance of the left beer can from the camera could control the drum volume while that of the right can could control the guitar volume. The variation in time of the left angle could control the tempo and the right angle could control the cow bell volume. The relative distance between the two cans could control a singer's volume. The smartphone (or other camera/computer device) can be manipulated by another player to modify the geometric properties of the targets as tracked by the Rich Color Transition Curves. Furthermore, the targets can be kept stationary, and the camera moved. The path and geometric properties, as well as the speed and acceleration of the Transition Curves in the sequence of video frames can be used to define a predefined macro or computer program or identify a specific database item to be used for with the macro or computer program.

In order to both reduce repetitive language in this description and clearly described the method, consistent numbering is used in FIGS. 9-19 such that the last two digits refer to items in the set of drawings is identified below and the other digits are the Figure numbers:

- 00—Objects colored in pale colors
- 01—Reddish colored human face
- 02—Object colored Rich Red
- 03—Object colored Rich Green
- 04—Object colored Rich Blue
- 05—Rich blue colored sweater
- 06—Object colored Rich Magenta
- 10—Target with Rich Red, then Blue, and then Green (also used in FIGS. 19, 20)
- 11—The Rich Red patch of Target 10
- 12—The Rich Green patch of Target 10
- 13—The Rich Blue patch of Target 10
- 16—Transition Curve between the Rich Blue and then Rich Green patches of Target 10
- 17—Transition Curve between the Rich Red and then Rich Blue patches of Target 10
- 20—Target with Rich Red then Green and then Blue
- 21—The Rich Red patch of Target 20
- 22—The Rich Green patch of Target 20
- 23—The Rich Blue patch of Target 20
- 25—Transition Curve between the Rich Red and then Rich Green patches of Target 20.
- 26—Rich Color Transition Curve between the Rich Green and then Rich Blue patches of Target 20
- 30—Target with Rich Red at angle then Green and then Blue
- 31—The Rich Red patch of Target 30
- 32—The Rich Green patch of Target 30
- 33—The Rich Blue patch of Target 30

18

35—Transition Curve between the Rich Red and then Rich Green patches of Target 30

36—Transition Curve between the Rich Green and then Rich Blue patches of Target 30

40—Target with Rich Red then Blue 41—The Rich Red patch of Target 40

43—The Rich Blue patch of Target 40

47—Transition Curve between the Rich Red and The Rich Blue patches of Target

57—Transition Curve between the Rich Red head 01 and the Rich Blue sweater

91—Noise

95—Target 20 center 96—Target 10 center 99—Non-Rich background.

FIGS. 10, 11, 12 and 13 show what would typically remain after respectively applying Rich Color Filters for Red, Green, Blue and Magenta to the original image shown in FIG. 9 assuming good cameras, lighting, and thresholds. If the bias for each color filter was set to a very high value, a binary image would result from each filter step. Invariably filtering in the real world introduces noise artifacts 1091 as seen in FIGS. 10-14.

FIG. 14 shows a composite monochrome image that is created from images shown in FIGS. 10, 11, and 12 such that Red pixels have a value of 1, Green 2, and Blue 3 (since the Rich Green and Rich Blue Binary Filtered images were multiplied by 2 and 3 respectively before added to the Rich Red Binary Filtered images) while everything else has a value of 0.

FIGS. 15, 16, and 17 show the results of passing the images from FIGS. 10, 11 and 12 through a Rich Color Transition Curve Finder for Red/Green Transition Curves, Red/Blue Transition Curves, and Green Blue Transition Curves respectively. The Transition Curves that separate patches of Rich Color from each other are then identified first as points that are then strung together into Transition Curves. These curves have additional characteristics of a start and end as well as a left side color and a right side color which is referred to as "Color Pairing". Color Pairing is very useful in identifying which Transition Curves belong to which targets. The geometry and color characteristics of each line are then stored and compared to a target database. The relative size, distance apart, number of Transition Curves, shape, orientation, color pairings, orientation of the color pairing, and other characteristics are used to identify a specific set of targets from other data. This should be a simple task since the remaining Transition Curves should be few compared to the original challenge of 1M or more of unassociated pixel data.

FIG. 18 shows four Transition Curves remaining after database identification leading to two targets. Two separate search techniques are presented in this one figure. The target on the lower left (original target 920 in FIG. 9) was found using the database to find a target with Transition Curves 1825 and 1826 that have the correct color pairing and relative angle to each other. The geometric center of these two Transition Curves is designated as 1895. The method used to find the second target (original 910 in FIG. 9) is the inline method that is described below in FIG. 20-23. FIGS. 22 and 23 show how the first two points of each of the two Transition Curves (1816, 1817) that come from target 910 are found. This method is well suited for hardware implementation.

FIG. 19 shows that the target is identified additional information can be gathered from the database or searched for in a region around the target or from the inline method itself. The size, orientation, and location of this region can be calculated from the geometric information gathered from the Transition Curves together with database search parameters for the

given target. If the database had physical information such as mass and moment, then forces and other physics can be computed from video images. Finally the target database can contain information on which hardware should be turned on or off or which software programs should be run and what parameters should be used to operate these programs.

By repeating the method with different threshold and camera parameters and then comparing the results with expected results of known targets for a fixed target and environment, it is possible to optimize the camera and threshold parameters.

FIGS. 20 and 21 show a flowchart that illustrates how the method can be implemented on a stream of pixels making it ideal for hardware implementation. Step 1 initializes the Search Process for an image of I rows and J columns. Start with $i=0$ $j=0$ (where i is the search-line number), and initialize the variables defining the targets and color characteristics. Step 2 Increments the search-line count $i=i+1$. Then test to see if i is outside of the image; if so end the process. Step 3 initializes Pixel information on Search-line i . Set the pixel number $j=0$ and purge the memory of "Line-Sets". We define a Line-Set as a set of three items of data defining a linear patch of data that has a common Target Color: the Color Indicator, the start pixel location, and the end location positions for color patch. Step 4 increments the pixel count $j=j+1$ and sets the Rich Color Indicator $S=0$. Then it tests and If $j>J$ (where J is the total number pixels in a search-line). If j is outside the image, return to Step 2. In Step 5 grab the image data for pixel j (N colors c_1, c_2, \dots, c_N from N color components). Then in Steps 6-9 test if all of the original components C_n are below a Black Threshold, T_B , or above a White Threshold, T_W . If true, then the Target Color Indicator is set to $S=-1$ or $S=-2$ respectively and then proceed to Step 12. If false, then proceed to Step 10 where to find the largest incoming color component C_H of the pixel j .

In Step 11 apply the Rich Color Filter to produce a flag setting representing the binary image value for the current pixel for each color component. A color made up of 1 or more color components that are much larger than the remaining color components as typically defined by the Rich Color Filter equation for a filter type "m" color distance equation. For a pixel located I, j the color distance equation is defined as follows in terms of ratio space color component values (c_1, c_2, \dots, c_N):

$$[A_m = \text{POS}(B_{1,m} * (f(c_1) - T_{1,m})) + \text{POS}(B_{2,m} * (f(c_2) - T_{2,m})) \dots + \text{POS}(B_{N,m} * (f(c_N) - T_{N,m}))]$$

If $A_m=0$ and the filter color m is a Target Color, set a Target Color Indicator, S , to the Target Color values; otherwise set S to a value indicating that it is a non-Target Color. $T_{n,m}$ is the threshold for color component n and color filter type m , $f(c_n)$ is a function of c_n , and the Bias $B_{n,m}$ (or "scale parameter") for color component n and color filter type m . The function POS (A) such that POS (A)=0 if A is less than or equal to 0, ELSE POS (A)=A. The function $f(c_n)$ can simply be any function that emphasizes Rich Color for example $f(c_n)=[E_1 * c_n]+E_2$ or it might be $f(c_n)=[E_1 * c_n * c_n]+[E_2 * c_n]+E_3$ for input parameters E_1, E_2, E_3 .

The Rich Color filter could be implemented in other ways. For example, each of the terms between the plus signs in the distance equation could be implemented with IF tests. Simi-

larly, a Rich Color Look-Up Table could have been used to produce similar results. The basic idea is the same; search for the presence of Rich Colors in the Rich Color Ratio Space.

FIG. 21 shows a method where the Rich Color Indicator, S , can be used along with two types of Data-Sets (Line-set and Point-Set) to find Rich Color Transition Curves that are then used to find objects and targets in images. Again, a line set is a set of consecutive pixels all having the same Target Color ID and containing three pieces of data: the start pixel location of the pixel cluster, the end location, and the Target Color ID. A Transition Point Data-Set (or just Point-Set) consists of four data sections: an ID section, a point location, a backward Line-Set, and a forward Line-Set. The ID section includes a Transition Point ID number, a Transition Curve ID number, a Back-link Transition Curve number, and a row/column indicator. The point location has a search-line number, and the Transition Point location (X) on the search-line. Last two sections are the Rich Color Line-Sets that flank the Transition Point. The location X is defined by a statistical measure of the distance between the nearest start and end locations of the two Line-Sets and any non-Target pixel colors that fall between these locations.

Steps 13 and 14 shows the image being processed along search-lines (rows, columns, etc.) initializing and updating data sets for new image and new search-line. If the new pixel's Rich Color Indicator, S , is the same as it was for the previous pixel then simply update current Line-Set with an end location= j . Otherwise go to step 16, FIG. 21, and create a new Line-Set with the location of the start and end both set to the new pixel location and the Target Color ID set to that of the new pixel. Also, when the Target Color ID changes, begin testing for a Rich Color Transition by checking if there is the start location of the previous Line-Set is within a tolerance distance of another Line-Set further back in the current Search-line such that both have different but non-zero Target IDs. Any Line-Set further away than a distance tolerance, T_{LD} , is usually purged from memory to minimize memory requirements. If a Line-Set is found within the distance tolerance, T_{LD} , then a Transition Point Data-Set is created or updated. Steps 19 and 20 in FIG. 21 show how a Point-Set of the current search-line can be used to link to Point-Set on a previous search-line to form a Transition Curve and then possibly link up to a previous Transition Curve on the current search-line to identify targets.

FIGS. 22 and 23 show a detailed example of how the method to find a possible Rich Color Transition Curve is applied to a set of pixels on a search-line. This example is similar to Target 910 that is seen in FIG. 9 with the threshold and bias values below:

a.	R	1.	.7	.7	-10000	10000	10000
b.	G	.7	1.	.7	10000	-10000	10000
c.	B	.7	.7	1.	10000	10000	-10000

Step 1—Input (R,G,B) -7 Step 2—Ratio Space (r, g, b) -7 Step 3—Rich Color Distance (A_r, A_g, A_b)
Step 4—Indicator Flags (q_r, q_g, q_b)-7 Step 5—set Target Color Indicator S

J	$C_1 = R$	$C_2 = G$	$C_3 = B$	$c_1 = r$	$c_2 = g$	$c_3 = b$	A_r	A_g	A_b	q_r	q_g	q_b
100	159	101	124	1	0.635	0.779	799	7447	4553	0	0	0
101	185	55	81	1	0.297	0.438	0	10027	8622	1	0	0
102	179	54	80	1	0.301	0.447	0	9983	5531	1	0	0

-continued

J	C ₁ = R	C ₂ = G	C ₃ = B	c ₁ = r	c ₂ = g	c ₃ = b	A _r	A _g	A _b	q _r	q _g	q _b
103	189	51	78	1	0.27	0.413	0	10302	5873	1	0	0
104	187	50	79	1	0.267	0.422	0	10326	5775	1	0	0
105	105	46	80	1	0.438	0.762	619	8619	2381	0	0	0
106	70	44	81	0.864	0.543	1	4358	9210	1642	0	0	0
107	48	43	81	0.593	0.531	1	7074	7691	0	0	0	1
108	50	44	83	0.53	0.53	1	6976	7699	0	0	0	1
109	48	43	84	0.512	0.512	1	7286	7881	0	0	0	1
110	51	46	87	0.529	0.529	1	7138	7713	0	0	0	1
111	47	46	85	0.553	0.541	1	7471	7588	0	0	0	1
112	103	108	99	0.954	1	0.917	5630	2537	6370	0	0	0

FIGS. 22 and 23 show how the above pixel data can be used to construct both Line-sets and Transition Point data-sets. A Red data set ending at pixel 104, FIG. 23, is within a tolerance distance of 4 of the Blue data set starting at pixel 107. It is irrelevant that there is a Line-set of non-Target colors between these two. All that matters is that the two are within tolerance of each other and of different but Target Colors. The Transition Point Y value is usually set to the current search-line number. A simple statistical measure of the Transition Point X value is an average of the end pixel number and start pixel number of the two Line-sets respectively. Another measure would be a weighted average of the pixel locations using the original red and blue component values to weight the band of pixels starting at 104 and ending at 107 as shown below:

$$X = [104*(187+79) + 105*(105+80) + 106*(70+81) + 107*(48+81)] / [187+79+105+80+70+81+81+81] \\ X = 105.2$$

The values above are used in FIGS. 22 and 23. In this example, it is assumed that scan-line 23 is the first line that crosses target 910. Thus, the first Transition Point is found at X=103.4 and a Transition Point-Set is created with a Point ID=1 and a Transition Curve ID=1 since these are the first Transition Point and Transition Curve of the image. But since there is no previous Point-Set on this search-line, the Back-link Curve ID=0. Later on the same search-line another Transition Curve is uncovered at X=111.4. This being the second Point found in the image and on a new curve the two IDs here are set to 2. The associated Back-link Curve ID is set to 1 since the Forward Line-Set in the last Point-Set is the same as the backward Line-Set for the new Point-Set (106, 3, 110). With no other Transition Points found in the image, the search-line moves on to search-line 24. Here in FIG. 23 a Transition Point is found at X=105.2 and a new Point-Set is created for Point ID=3. Then look for Point-Sets on the previous search-line that is located within a tolerance of Point ID=3 with the same color pairing. Since Point ID=1 passes both tests, the Curve ID of the new Point-Set is set equal to that of Point ID=1 (i.e. the new Curve ID is set to 1). However, there is no Point-Set to the left of the new Point-Set and so the Back-link Curve ID is set to zero. Later on the same search-line another Transition Curve is uncovered at X=112.4. This being the fourth Point found in the image, the Point ID=4. Searching the previous search line for a Point-Set within tolerance and having the same color pairing leads to setting the new Curve ID to 2. The associated Back-link Curve ID is set to 1 since the Forward Line-Set in the last Point-Set is the same as the backward Line-Set for the new Point-Set (107, 3, 111).

The resulting curves together with their flanking Rich Color patches can be compared to target properties in a target database. By creating a set of Transition Curves, the original problem is reduced to one of comparing a database of target properties to a small set of curves with simple geometric and

color properties. The absolute and relative size, and orientation as well as shape and color can be used to detect, identify, locate and orient (in 2D or 3D space), and characterize a set of objects or targets. A sequence of images together with associated database properties can be used to determine physical properties of the objects to which the targets belong. A sequence of video images can be used to define physical properties such as velocity, momentum, acceleration, etc. and further filter the data by enhancing image quality (averaging target data) for targets that have matching characteristics. These characteristics should allow for further tracking refinements and rejection of potential targets that cannot physically change properties radically as seen from frame to frame.

Once a target has been identified along with its geometric and other properties, these can be used with a database and a lookup table to manipulate computer interfaces, computer programs/macros, or devices. Both hardware and software implementations are disclosed. A smart camera with this method implemented in hardware could analyze each frame while the next frame is being acquired and then transmit (wired or wireless) a small data set of Transition Curves only if a target is detected in the frame. This method requires only a small amount of buffer memory and a very few multiples, adds, if tests and bit manipulation per pixel processed. The speed and flexibility of the method make it possible to use a variety of targets at the same time. If targets with unique color and geometry are used, then each target that is found can trigger a unique action based on their absolute or relative location and orientation together with their associated database properties.

FIG. 24 shows how a Sequence of Rich Color Transition Curve Targets can be implemented on a tape 2400. The side edges of the tape have measurement markings (2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, and 2471) like a tape measure. The targets (2401, 2402, 2403, 2404, and 2405) in this example each have four Rich Color stripes that are separated by any one of five distinct interval lengths measured from the centroids (2441, 2442, 2443, 2444, and 2445) of each target. It is helpful to mark the tape with special lines (2421, 2422, 2423, 2424, and 2425) that both passes through each target centroid and is perpendicular to the line through the target centroids. The targets in this example are limited to five distinct rotation angles in degrees (-20, -10, 0, 10, 20) as measured from the line between target centroids. These targets have one of three distinct sizes (small, medium, large). Each target is shown with an identification number (2431, 2432, 2433, 2434, and 2435) to its left. Even if the number of Rich Colors are limited to three, Red, Green, and Blue, there are 120 unique color patterns of four stripes for each of these five targets. This allows over a billion unique target combinations that are part of such a sequence of five Rich Colored targets to be identified. This is an incredible result for such a few set of parameters. A roll of Rich Color

Sequence Tape could have unique sets of five targets but a set of three targets is more likely since there is unlikely to be a need for a billion unique combinations of target sequences. FIG. 24 also shows the optional use of small "satellite" targets (2451, 2452, and 2453) in the neighborhood of a central sequence target. These provide an additional way to differentiate targets.

The complete implementation of the Rich Color Target Sequence would include both the sequence and its corresponding database. An example of such a database is shown in FIG. 25. The database can be made available via the internet or other data storage method. The database contents would include the relative location, and orientation of each target and the interval between each target along with their length and width. Further, the database would include the target shape and Rich Color patterns.

The example of a Rich Color Target Sequence in space is shown in the table in FIG. 25 where a set of targets on a ribbon have unique characteristics. Target2 and Target4 in this sequence are identical in shape, color, size, and orientation but the intervals between their neighbors on the left and right differ greatly. The sequence of target shapes, color patterns, area, orientations, and separations can be stored in a small database associated with the ribbon/tape. The same sequence in FIG. 25 could apply to define a Rich Color Target Sequence in time if the interval between targets was in fractions of a second not inches and the sequence was played like a video of varying display rates on a single LED display.

A Rich Color Target Sequence (RCTS) tape, together with its database, has the added advantage that makes it simple and economically possible to setup a variety of applications with the same hardware and methods. Examples of such applications are movie making, physical therapy or health applications, teaching how to use handheld tools (from surgical to carpentry), scientific measurement, CAD interfaces, or gaming to name a few. A preprinted ribbon that includes Rich Colored Targets could be tacked onto walls, equipment, furniture, or other items or sown into clothes or armbands. Just cut off the required length of ribbon, identify the target numbers on the ribbon for the start and end of the ribbon segment used, load the capture computer with the database information for this ribbon segment and go. Often current machine vision applications have specific codes written for specific applications. This is expensive and the loss of generality leads to small markets with buggy application code. The use of RCTS tape should lead to much more robust applications.

If inexpensive passive targets were used outdoors without specialized lighting or targets, the number of pixels captured by ordinary cameras would be staggering. Imagine the number of pixels that have to be analyzed if the motion tracking was done on a city street. In order that each actor is captured by two or more camera to capture stereo pairs to triangulate 3D location and orientation over a large real world scene, large number of cameras would have to be located at overlapping locations and orientations to get good resolution and avoid occlusion as one actor gets in front of one camera or another. The higher the resolution of each camera, the fewer cameras would be needed; but the total pixel data remains roughly the same. For professional applications, hundreds of cameras might be used to capture the scene from multiple angles and points of view.

The Rich Color Transition Curve tracking method dramatically reduces the volume of data that needs to be collected from any camera frame. In this example, consider a video camera with a frame resolution of 4000×4000 pixels capturing ten targets such that each target is made up of three Rich Color stripes and covers an area of 10×10 pixels. If each pixel

took up three Bytes of memory (0-255 for each color), a frame of RAW data would take up 48,000,000 Bytes. Now let's consider a hardware implementation of the Rich Color Transition Curve method applied in-line with the camera capture. Assuming that the transition points are stored as long integers and each curve had a long integer to define the color pair involved and the right/left orientation of the pair, then the Rich Color Transition Curve tracking would require only 10×(10+2)×2×2 Bytes or 4800 Bytes, thus, reducing storage by a factor of 10,000. By converting each target to a 3D vector and a colored identification data set, the data that must be transmitted and stored for post-processing is roughly 50 Bytes per target per frame per camera. This is roughly 500 Bytes in this example reducing the data by a factor of 100,000.

If this camera were operating at 128 frames a second to capture an action sequence for 10 minutes, the RAW data file would be 3.686 Terabytes. The Rich Color Transition Curve vector file for this would be roughly 40 Megabytes or less than 1/3 the size of a RAW file from a single high resolution photos of the Nokia Lumia cell phone. This is small enough for 100 such cameras to easily transfer their capture data files to a cloud storage facility in real time. Likewise the small number of arithmetic operations required for the Rich Color Transition Curve method means that the data file could also be created in real time. In the future the number of pixels per frame will skyrocket. Since the Rich Color Transition Curve method is roughly linear to pixel count of the camera, this should not be problematic.

For movie making, it is always more desirable to have small inconspicuous targets. The smaller the target, the less curvature that can be detected and the less variation between Rich Color Transition Curves that can be distinguished. Ideally higher resolution cameras of the future will clear this up. But for now, the most common targets for our actors will consist of 2 or 3 simple straight uniform stripes. These fit the bill for small output files and that are fast to compute. However cameras today generally have poor color resolution compared to what is expected over the next decade. This means, for example, that while a green pixel sensor may have peak sensitivity in the green part of the spectrum, it is sensing a lot of photons from the Red or Blue part of the spectrum as well. Until color resolution improves and until new sensors are added to the typical three (RGB) that make up current camera data, most Rich Color Transition Curve implementations can only use three Rich Colors. For example, four unique two stripe targets and twelve unique three stripe targets may be provided. But for our street scene we will need hundreds of targets.

This is still better than the situation for some passive target systems that use retro reflectors where all targets look alike. They get along by first identifying each target and then track it in each succeeding frame by finding the target that most closely matches the location, orientation, and vector velocity on the last frame. There can be a tedious startup identification and additional post processing work if a target is occluded at some point. The sequences provide simple identification even in cluttered environments with multiple cameras and large number of actors and objects to be tracked.

This is where the method that will be referred to as the Rich Color Target Sequence (or RCTS) comes in to play. This method defines far more uniquely identified targets. It is more automatic. This method delivers far better overall 2D and 3D location and orientation accuracy. Also calibration and scene stitching are easier. All this comes from the ability to identify a large number of unique sets of targets with accurate relative geometry using the corresponding RCTS database. Other

25

targets can be identified located and oriented based on their proximity to one or more RCTS in the frame. Targets in the RCTS can be identified by their relationship to the members of the sequence.

Ribbon sequence allows one to easily attach large numbers of targets on the periphery and interior of the capture volume and on actors and objects. The sequence allows both accurate location and orientation and identification. Use one set of target patterns for references and another set for actors and objects. For example the actors and objects could use Bull's Eye targets with two or three Rich Colors while the ribbons used in the background for reference could use four Rich

Another advantage of this method is that these targets can be located and removed from any frame, using automated image editing to replace the target with a predefined image.

A restricted version of the Rich Color Transition method can be used when lighting is well controlled over time, roughly uniform, and the cameras involved have excellent low light capability. To understand this, consider an image composed of $k=1, K$ pixels each having N color components C_i and divided into Rich and non-Rich colors by a color component thresholds T_i . Then the sum of the color components for pixel k can be written as " $a(k)$ " such that

$$a(k)=C_1+C_2 \dots CN$$

If the illumination is fairly uniform and constant over time, then $a(k)$ can be approximated by a constant " A ". When applying thresholds to Rich Color Transition Curve methods for this restricted class of problems, there are times that the standard thresholding for a pixel k

$$T_i < [C_i/a(k)]$$

can be approximated by

$$[T_i * A] < C_i.$$

FIG. 26 shows a set of RCTS on the walls of a room being used for a motion capture application. Sequence 2611 is painted on the side wall. Sequence 2612 is tacked onto the back wall 2602. Sequence 2613 is taped vertically onto the back wall while sequence 2614 is displayed vertically on an LED screen. Three smart cameras (2631, 2632, and 2633) are attached to the walls next to Rich Color Target Sequences. A stereo pair of cameras (2634, 2635) is positioned looking at the back wall.

If the room is ringed with cameras and target sequences, every part of the room can be seen by cameras that in turn can see RCTS that have cameras attached at known points. Each camera can see at least five targets as well as other information for identification and geometry. The measuring tape markings on the ribbon or tape can be entered into the database to further define the location of wall cameras or sequence start/stop target locations relative to corners in a room. Cameras 2634 and 2635 are capturing multiple target sequences traversing the image horizontally as well as vertically to produce accurate calibration. Since they are seeing many or the same targets, standard stereo triangulation can be used to calibrate the 3D capture space. Cameras that can see the location of other cameras next to target sequences can use this information to calibrate those cameras. Multiple cameras ringing a capture volume provides redundancy in case of occlusion and removes pixel truncation errors.

In FIG. 26 a box 2651 is attached to a ribbon used for sequence display 2641.

The ribbon comes in the box 2651 and is drawn out for the required length with the unused portion of the ribbon remaining stored in the box. The ribbon is composed of a display technology such as an OLED display on a flexible transparent

26

material such as a plastic film or Corning WILLOW GLASS. The box 2651 also contains a microprocessor 2652 with internet communication 2653 capability that can take color pattern data files from the cloud and display them on the ribbon's OLED display. An attachment tab 2642 of the ribbon is attached by thumb tacks, nails, adhesive or the like to a wall or other surface. This OLED ribbon can display a Rich Color Target Sequence for calibration and again display the full sequence or a subset as required by a machine vision application. When no machine vision applications are in use, the ribbon can display something pleasing such as an ornamental design or simply the wall color. For some applications the targets can be displayed as the same targets as were used in calibration except that their color can be a slight variation from the background color of the ribbon. In this case only the region of the image where the ribbon and the target used to be are searched. Any pixels in this region with color components that differ from the ribbon background color by more than an input threshold value will be mapped into the corresponding Rich Color component of the original target. This would allow us to use the Rich Color methods to be used in a less conspicuous manner.

It is anticipated that very small smart cameras (cameras with communication and computers) will hit the market in a few years that are basically smart phones without the phone or display shrunk down to the size of an iPod Shuffle (roughly 1.5x1.5x0.3 inches). Ideally such a camera should have a built in hardware version of Rich Color Transition Curve tracking. Such cameras could be placed around a virtual reality rooms used for gaming, teaching, health, or other applications. Simple motion capture setups using these small low-cost cameras will use RCTS ribbons to both redundantly define the capture volume and to define the camera locations and orientations. By attaching the cameras to these ribbons, the locations and orientations of the complete camera set can be computed using overlapping camera images together with the geometric information associated this target sequence to iteratively calibrate the 3D geometry of the room or capture volume. Such rooms could have RCTS Ribbons around the walls, floors, and ceilings. By placing the targets next to sequence targets, the location orientation and identity of each camera can be determined from surrounding cameras. The same can be said to define the location, orientation, and identity of RCTS ribbons. RCTS can be displayed on an LED display also. Thus setup and calibration of such a capture volume could become automatic.

The more reference targets the more accurate the calibration. The more frames that are averaged together, the more accurate the calibration. A panorama can be constructed from an overlapping set of images that share a common tape of RCTS using Transition Curve targets. If one were setting up a movie scene with motion capture on a city street one might ring the capture space with horizontal strips of Rich Color Target Sequence tapes and occasional vertical strips.

FIG. 27 shows a flat surface (a wall, a floor, a table top, a kitchen counter, a driveway, etc.) with a set of RCTS attached to this surface at reference locations (corners, edges) and/or next to objects of interest. For this discussion we assume that the surface is a wall. Two Rich Color Target Sequence Tapes (2701 and 2702) define the top and left side of the wall. In this example a picture 2710 is hung on the wall. The targets of the two tapes together with their corresponding database information are enough to calibrate any image which includes the tapes using standard calibration methods.

The camera 2760 found on the back of the tablet 2750 is used to take a photo of the picture on the wall together with the tapes. The image is displayed on the touch screen display

2770 of the tablet 2750. The two tape sequences on the top and side of the wall that are captured and displayed in the photo image as 2731 and 2732 are used to calibrate the image. By touching any location on the screen such as the lower corner of the picture frame, one can make measurements on the image such as the distance to the upper left corner of the wall. Much more accurate measurements can be made by attaching additional sequences to points of interest. By adding additional sequence tapes, 2703 and 2704, starting at the lower left and upper right corners of the picture frame 2742 and we can accurately compute the point where the two lines would meet using the pixel data from all of the targets that make up the target sequences (2703, 2704) can be accurately computed.

FIG. 28 shows a simple rectangular room with three walls (2800, 2801, 2802) that meet at right angles and are covered by a ceiling 2805. Throughout this application a "room" is used in many drawing figures for the simplicity making a drawing; but really this could apply to most any space bounded by or containing planar surfaces. FIG. 28 shows a room that has a multitude of target sequences (2810, 2811, 2812, 2813, 2814, 2815, 2816, 2817, 2818, 2819, 2820, 2821, 2822, 2823, 2824, 2825, 2826, and 2827) and cameras (2831, 2832, 2833, 2834, 2835, 2836, 2837, 2838, and 2839) on the walls.

One common way to calibrate the cameras and the room as a whole is to use a 2D array of dots or squares displayed on a movable surface such as a sheet of cardboard 2890. In this example, the camera's intrinsic matrix is computed in the calibration process. This movable surface could also be a sheet of paper or a tablet or large yet thin OLED sheet with a microcomputer running the display. The tablet or OLED might be a better choice since the size of the array could be easily changed to match the camera zoom setting. Capture one image of the array precisely placed next to a location marker 2891 on a long sequence tape 2832 that stretches completely across a wall horizontally. Then capture one or more images on the same camera with the array moved and rotated in 3D in the field of view of the camera. Repeat this for all cameras in the room. This method both defines the intrinsic matrix for the camera and defines a 3D coordinate system for the camera field of view that is attached to the sequence tape on the wall. Now use an optimization routine to tie all of these camera coordinate systems into a single 3D coordinate system that represents the whole room. Triangulation with stereo pairs of cameras looking at any uniquely identified target can be used to better define the 3D position within the world (room) coordinate system.

The following describes another more automated calibration procedure for a room such as that shown in FIG. 28. This assumes that a camera database exists for each camera that contains the intrinsic matrix used to take out lens distortion, pixel aspect ratio, and other such problems. Use the intrinsic matrix to matrix multiplication of the pixel location of each target centroid in the camera plane and then solve for the position in world coordinates that corresponds to the target position in the camera planes of multiple cameras using inverse rotation and translation matrices computed by the calibration method described next.

This also assumes that the geometry of the room is provided in a 3D CAD (Computer Aided Design) database. Consider a camera aimed so that its field of view contains the corner formed by the wall 2801 and the wall 2802 and portions of at least two target sequences for each wall. In this example use sequences (2815 2825) for wall 2801 and sequences 2816 and 2817 for wall 2802. Further, let the field of view capture at least three targets per sequence. The targets

and 3D CAD geometry are enough to make each wall serve the same purpose as the planar sheet 2890 shifted to two positions as described in the paragraph above. Again, a more accurate 3D data would be obtained if a second camera could see the same corner and triangulation was used. Place smart camera on the walls with overlapping field of views each seeing sequence tapes that other cameras can see. In general this image data together with the 3D CAD data can be used to stitch together multiple camera images and solve the 3D position of targets within the room.

The fact that a target sequence is typically formed by targets whose centroids fall on the centerline of a tape can be used to greatly enhance accuracy of any results. After multiplying the vector made up of target centroid locations in the camera image by the intrinsic matrix found in the camera calibration, the centroids must fall on a line except for image error. This is result that falls out of the affine geometry that defines the relation between the camera plane and any object plane in the field of view.

FIG. 29 shows a sequence of circular Bulls Eye targets laid out on the centerline of the tape 2900. Each of these Bulls Eye targets (2910, 2911, 2912, 2913, 2914, and 2915) is constructed with three concentric circles each of which is composed of a Rich Color. Alone these targets have no preferred directionality in the plane of the target. Previously it was described how image data can be traversed on one or more times along different scan line directions. However, it is faster on today's computers to only traverse an image with horizontal scan lines. Thus, targets are most valuable if they tilt only moderately in the target plane (plus or minus 30 degrees) such that transition curves are mostly vertical. In cases where the targets can appear at any angle in an image circular Bulls Eye targets are preferred. The sequence database together with the image capture data allows one to construct very accurate angle data and pretty good location data. The ability to differentiate one sequence target from another can be enhanced by placing a small satellite target close to it at some specific angle. In this example the main targets of the sequence (2910, 2911, 2912, and 2913) are laid out on the centerline can be identified as having four transition curves (i.e., five stripes) and the ID modifiers or satellite targets have only two transition curves (i.e., three stripes). The number of satellite targets, their color pattern, and their angle relative to the centerline provide additional methods to distinguish a target on a RCTS tape.

FIG. 30A shows how a circular target would appear as a set of striped targets to a horizontal scan line implementation of the Rich Color Transition Curve method. The target 3000 is composed of three Rich Color circles—Red 3001, Green 3002, and Blue 3003. A set of horizontal scan lines are shown passing through the center of a Bulls Eye target. In this example four Rich Color Transition Curves are created which serve to detect, identify, locate, and orient the target. A convention that scan lines pass from left to right is used. Thus, in FIG. 30B transition curve 3021 transitions Blue to Green. The transition curve 3022 transitions Green to Red. The transition curve 3023 transitions Red to Green. The transition curve 3024 transitions Green to Blue.

FIG. 31 shows a set of sequence tapes (3110, 3111, 3112, 3113, 3114, 3115, 3116, 3117, 3118) attached to an actor's body 3100 and clothes viewed by a video camera 3150. As described previously in the case of striped targets, unique sequences dramatically increase the number of unique identifiers and also greatly increase the accuracy of position and orientation for measurements.

Rich color targets do not have to be printed on a stark white tape. They can be buried in a colorful scene that consists of

non-rich colors except for the targets. Your eye naturally averages out the colors and makes the targets hard to see unless you without the use of the Rich Color Transition Curve method. The reverse is true for the Rich Color Transition Curve method for which white or black or pale colors and any colors that are not designated Rich Colors that will produce Rich Color Transition Curve separated appropriately by designated ratios will appear to be invisible.

In FIG. 32 targets are hidden in an artistic display so that they are not overwhelming to the human eye. The center of flowers 3201, 3202, and 3203 are made up of 2 Rich Colors and form a Rich Color Target Sequence of 3 targets. The rest of the flowers are limited to any pale color with color components below the Rich Color component thresholds, black, or white. If target colors are limited to Rich Red, Rich Green, and Rich Blue, then the painting could also use Rich Yellow, Rich Magenta, and Rich Cyan. It could even use one circle with Rich Red, Rich Green, or Rich Blue and have the other circle with a non-Rich color since no Rich Color Transition Curve would result

So that one can imbed a sequence of Rich Color Targets in an artistic display such as a tape with a floral pattern as shown in FIG. 32 as flower targets 3201, 3202, and 3203. with Rich Color Target Components of Rich Red and Rich Blue on the outer and inner 3221, Rich Blue 3222, Rich Green 3223, Rich Blue 3224, Rich Red 3225, Rich Green 3226, and Rich Red 3227.

FIG. 33 illustrates a shower stall that uses smart cameras (3321, 3322, and 3323) that are a hardware implementation of the Rich Color Transition Curve. The person taking a shower (3301), hereafter called the bather, in FIG. 32, is shown as a stick figure in a shower stall (3300) with RCTS (3311, 3312, and 3313) implemented in bathroom tile on the shower walls. A camera with hardware implementation of the Rich Color Transition Curve method that includes a microcomputer and communication capable to sending and receiving cloud data can be manufactured so that the only output possible are data sets defining each target detected such as the target's centroid, orientation, color pattern, and size but no images are output. Such a device will be referred to as a Secure Smart Rich Color Camera or SSRCC. Without any output image, the SSRCC is as innocuous as a light switch preserving the bather's modesty. Data from the SSRCC (3321, 3322, and 3323) can be sent wirelessly to computer 3371 without the fear that someone with malicious intent might intercept camera images of the bather and publish it on the internet. Since no image data can be output, such efforts would be thwarted. The use of normal cameras in home settings potentially places the residents at risk.

In this example, SSRCC 3321, 3322, and 3323 capture Rich Color Transition Curve data sets of the Target Sequences in the tiles on the shower stall wall and send the data sets wirelessly to a computer 3371. The computer compares them to the stored data set when nobody was in the shower stall. The transition curves that are missing compared to the stored file represents those that are occluded by the body of the bather. The cameras from each wall can indicate where the bather is and which shower head must be turned on or off so that the bather is covered with water but otherwise water is not wasted on empty space. The computer 3371 takes in the camera data and computes which shower head (3341, 3342, 3343, 3344, 3345, 3346, and 3347) must be turned on and which must be turned off. It then sends a wireless message to the microcontroller (3351, 3352, 3353, 3354, 3355, 3356, and 3357) that controls each motor driven valve that allows water

to pass from the water source pipe 3331 the corresponding shower heads (3341, 3342, 3343, 3344, 3345, 3346, and 3347).

FIG. 34 is a block diagram showing how machine vision applications can repetitively use prepackaged vision hardware and software components to solve problems with minimal development. In general machine vision applications are difficult to develop, and time consuming to debug. Only by using experts in the field does one stand a chance of solving even simple problems. When done does manage to get a solution, any minor change to the hardware, the environment, or the lighting can make the whole setup fail. A large number of machine vision applications can be developed with minimal code development by use of rich colored targets and sequences of Rich Colored Targets as shown in the block diagram. As computing devices and cameras get to be small inexpensive devices that can effortlessly send data packets back and forth to each other or to interim cloud storage devices, we can create a set of hardware/software modules that are reusable for many different applications.

The current software trend is to avoid using markers in machine vision applications. But the world is infinitely complex. Typical marker-less application code is written for a limited environment and set of objects to be analyzed. Using Rich Colored Transition Curve targets, we can remove many environmental problems such as variable lighting. We also can remove the need for specialized illuminators, or projection devices. These are passive targets that can be nearly as cheap as paper. One can use as many as necessary and by using them in special sequences as described above to uniquely identify them; And since one can uniquely identify each of hundreds or even thousands of targets, we can easily compare their locations and orientations from multiple cameras. Averaging this data would allow much more accurate location and orientation. Further, 3D locations and orientations can be established as well panoramic data sets.

FIG. 34 illustrates a method to make an ordinary room responsive to the occupants using machine vision technology. This system is designed for quick installation and easy modification. New applications can be added in the same way that one adds new apps to a tablet or cell phone. This system relies on six basic elements:

- (1) A cloud storage/retrieval technology that allows software and hardware elements to store and retrieve data between elements and where apps can be stored along with operational parameters.
- (2) Smart video camera that consist of computational power capable processing image data with Rich Color Transition Curve methods and handling standard I/O communication with a cloud
- (3) Rich Color Transition Curve targets and Rich Color Target Sequences that can be simply, inexpensively, and prolifically placed around the room as the applications dictate.
- (4) Computer software capable of using Rich Color target data together with application specific insert able routines to accomplish specific applications.
- (5) Computing devices to process the software in (4) and to handle standard I/O communication with a cloud.
- (6) Computer hardware capable of handling standard I/O communication with a cloud and utilizing the results of a given application. The devices here will probably be typical smart phones or tablets.

One can envision an interactive room similar to the example of the shower stall surrounded by Rich Color Target Sequence tapes and filled with hotspot locations using such tapes. The room has smart cameras at various locations and

31

angles such that the volume of the room is covered redundantly. The redundancy allows for 3D calibration and fully stitched together 3D space. We use "room" as a proxy for any capture volume such as a factory, warehouse, studio, living room, street scene, etc.

FIG. 34 shows how a vision application can be divided into three process levels each handled by a set of autonomous devices. Each process level is connected to the following level by intermediate data storage preferably a cloud. For simplicity from this point forward, the term "cloud" is used when more generally one could use intermediate data storage of many types. The control parameters (including when to start and stop) that define how the devices in each process level will be operated are stored in a cloud file 3410. The parameters are sent to cloud file 3410 from any smart phone, tablet, or computing device connected to the internet. Every device on Process Level 1-3 an ID and an on/off parameter associated with this ID. Likewise every application "app" has an ID and an on/off parameter associated with its ID. Similarly, to turn on or off or change any application, a user sends a parameter file to the appropriate cloud file. Each device (3431, 3432, 3433, 3434, 3341, 3442, 3443, 3451, 3452, 3453, and 3454) polls the cloud layer above their level searching parameter packets for identifiers of data packets to operate on or control parameters to initiate, modify, or stop their processing as dictated by these parameters. The first process level utilizes a set of smart cameras (3431, 3432, 3433, and 3434). A smart camera means a video camera with an attached computing device and communication capable of sending and receiving cloud data. The smart camera will apply the Rich Color methods to camera data. One implementation is the hardware implementation of the Rich Color Transition Curve method described above. Each frame of the smart camera is processed using the Rich Color Transition Curves method and then sent as data sets to a cloud storage file 3411. The Transition Curve data can be processed immediately and sent as target centroid, color pattern, orientation, and size or sent via the cloud as transition curve data and processed later into target data by application computers (3441, 3442, and 3443) of the second process level. Included in each camera data set is the ID of the camera, the time, date, and frame number. The format of these data sets is consistent so that any application operating on a computer in Level 2 can read and process the incoming data in the same fashion. Process Level 2 has a set of computers (3441, 3442, and 3443) that operates on either target data and solves for a complete application or operate on a subset of the application where it reads the input of one computational subset and outputs the result of that subset to be used by the next subset until all of the computational steps have been completed. The final results are sent to a cloud file to be handled by the third and final process level. Process Level 3 displays, sends the result to storage files, uses results to trigger an action or otherwise uses the result data. The most common devices of Process Level 3 are Internet TVs 3455, or mobile computing devices (tablets and smart phones) for task such as displaying results 3451, send data 3452, turning on equipment 3453 or playing music 3454.

It is only Process Level 2 that has any specific application code. If the application software is written as a sequence of routines each of which reads an input data set and outputs a data set for the next step, the application specific code can be downloaded to the Level 2 computers when the control parameters are initialized or changes and then downloaded.

The steps of Process Level 2 applications will start with using the color pattern and neighboring target data to identify the sequence that a target is part of and from this the target its location and orientation. The sequence database information

32

for the target can then be used together with the cameras calibration data to transform the target location and orientation into absolute coordinate system. Note that the first frame data of each camera will be used to establish the calibration data. If the camera is not stationary relative to a set of background RCTS, then a subset of the calibration must be performed for each frame. Once all of the targets in the video frame are detected, identified, located, oriented and transformed to the calibrated coordinate system, the application code can be used to process this information and produce a result data file that is sent to the cloud. The result data file will include all of the information that will be needed to use the result such as which application this belongs to, the time, date, cameras used, etc. A target rigidly attached to a rigid body with a geometric definition available can be used to locate or orient any part of that body in time and space using multiple video cameras as described herein. The hard thing in machine vision applications is reliably coming up with this information. Typically, the rest of the application (making use of this information) is the easy part.

To better understand, consider the examples shown in FIG. 27 and FIG. 33. In the example of FIG. 27 we use the memory of the tablet 2750 is used instead of true cloud memory. Also all of the devices for Process Levels 1 through 3 are that same tablet 2750. A simple interface step or app sets the on/off parameter for this example app to on. The software of Process Level 1 polling the memory file sees the app ID has an "on" value. It then loads this value and the setup data for this app into Process Level 1. The camera of the tablet 2760 then captures 1 image as setup parameters dictate. The image is processed to capture Rich Color target data using the Rich Color Transition Curve method. Finally, the memory file (or the cloud storage 3411) is sent the target data along with its app ID and camera ID and its polling event indicator set to 1. Level 2 has a program the constantly polls for the event indicator set to 1. Upon finding it, this event indicator is reset to 0 and a set of repeatable machine vision tasks are performed using the target data. First the target data is used to detect, identify, locate and orient every target using target data together with target sequence data. Then this information is passed into a calibration function. If more than one cameras were part of this application, then the pairs that could be used for stereo image data would be used for 3D calibration and the total assembly would be used to develop a panorama calibration. Further the results would be used to optimize the accuracy of the calibration. Finally all sequence marker locations and sequence angles are computed in the calibrated coordinates.

Only at this point is it time to use specific application code. This app makes measurements between two corners of an object identified by markers on two Rich Color Target Sequence tapes. The app asks the user to touch near two corners of objects in the display image. The app then searches for the targets nearest each of the two touches. Then the app finds the tape sequence that each of these targets is part of and then finds the sequence mark nearest the touch positions. Finally using the calibration, the coordinates of the two sequence marks is computed and placed in a memory file. Process 3 identifies that a new data file from Process Level 2 is available. The software of this level reads the point data, recognizes the data as defining a line which it displays along with the value of its length.

In the example of FIG. 33 true cloud memory is used, one camera one each wall of the shower stall, one computer 3371 to process Level 2, and 7 microprocessors 3351-3357 to control the valves for seven shower heads (3341, 3342, 3343, 3344, 3345, 3346, and 3347). A simple interface step or app

sets the on/off parameter for this example app to on. The software of Level 1 polling the memory file sees the app ID has an "on" value. It then loads this value and the setup data for this app into level 1. The cameras on the walls begin to capture frame images as setup parameters dictate. These images are processed to capture Rich Color target data using the Rich Color Transition Curve method. Finally, the memory file (or the cloud storage 3411) is sent the target data along with its app ID and camera ID and its polling event indicator set to 1. Level 2 has a program the constantly for the event indicator set to 1. Upon finding it, this event indicator is reset to 0 and a set of repeatable machine vision tasks are performed using the target data. First the target data is used to detect, identify, locate and orient every target using target data together with target sequence data. Then this information is passed into a calibration function. If more than one cameras were part of this application, then the pairs that could be used for stereo image data would be used for 3D calibration and the total assembly would be used to develop a panorama calibration. Further the results would be used to optimize the accuracy of the calibration. Finally all sequence marker locations and sequence angles are computed in the calibrated coordinates.

Only at this point is it time to use specific application code. This app finds targets that are covered by the bather's body. By comparing which targets are visible in the calibration step to the targets no longer visible in the current frame from each camera, we can calculate where the bather is standing. Then shower heads needed to cover the bather with water are identified to be turned on and the rest are identified to be turned off. Finally this shower head on/off data is placed in a memory (cloud) file. The microprocessors in Process Level 3 each identifies that a new data file from Process Level 2 is available. Each microprocessor acts to turn their valve to a new on/off position if the data file dictates.

To summarize:

- 1) This method uses inexpensive passive Rich Color Transition Curve targets by themselves or in Sequences such as tapes described above. The Rich Color methods allow us to use ordinary lighting and automatically ignore all but the targets in the room. From this point on, the only camera data that is used comes from target data packets greatly simplifying machine vision solutions.
- 2) Since the targets and sequence tapes are inexpensive, targets are used liberally so that a significant number of targets are seen by two or more other cameras allowing for 3D calibration around the whole room.
- 3) Using the Rich Color methods described in this application, it is possible to identify any target that is not blocked by a person or object.
- 4) Large numbers of cameras can be used to cover all angles and focus needed for the application. If the cameras are placed next to target sequences, it is possible to use the associated sequence database together with target data from multiple cameras to calibrate the full room.
- 5) The Rich Color methods are essential for cloud usage since only when the data is reduced to a small data set that can be transferred in real time.
- 6) Since most residents value their privacy, home applications that use cameras connected to the Internet are likely to appear risky. It is best to use the hardware implementation of the Rich Color Transition Curve as discussed in the example of the shower stall which will show nothing but a few lines of target transition curves.
- 7) The applications considered in this method all use targets attached to rigid bodies. Knowing a target's 3D orientation and orientation allow a computation of the same for any part of the rigid body.

- 8) Everything up to this point is the same for any application. Process Level 2 reads the target data in a standardized format independent of the application. Then each target ID, location and orientation needed for the specific application is sent to the application code and a result is calculated and sent to the results cloud file.
- 9) In Process Level 3 each device that uses the cloud results of Level 2. Each device in Process Level 3 constantly poles the cloud until the result file that it operates on has changed and is now available. Then this device uses the data as its setup parameters dictate.
- 10) Note that only a portion of the code in Process Level 2 is unique to an application. Also, most of the code in Process Level 3 is composed of reusable code.

Description Of The Method

U.S. patent application Ser. No. 12/107,092 discloses a means of filtering out all but Rich Colors of interest within an image under inspection. The resulting Rich Colored patches can be used to identify and track targets or Rich Colored objects.

In this patent application, the Rich Color methodology is extended and used to replace tracking methods like blob analysis. This is a robust method using a computer or embedded computer chip or specialized camera circuitry to rapidly find the location of a target within an image by searching for Rich Color patches adjacent to each other within a tolerance. This process can be performed in one process as it sweeps across the image one search-line at a time. An image with N color components can have up to N factorial (N!) Rich Color types. In the case of a RGB image (N=3), six possible color types can be identified for a pixel (Red, Green, Blue, Cyan, Magenta, Yellow). The sub-set of color types that are searched for are called Rich Colored Target Colors or just Target Colors. If the color type of the pixel is not a Target Color, the color indicator of the pixel is identified as a "non-Target Color" type. This method searches for Target Color patches different color type that are located adjacent to each other within an input tolerance. The method creates sets of adjacent pixels that have a common color type (including the color type of "non-target color"). The image is processed along search-lines (rows, columns, diagonal lines). Anytime on the same search-line that the end of one such pixel set is located within a tolerance of the beginning of another pixel set from the same search-line and both have different Target Color types, then a Transition Point can be defined. The Transition Point is located along the search-line at a weighted statistical location between the end and the start locations of these two pixel sets. These Transition Points can be strung together forming Transition Curves whose geometric and color properties and proximity to other Curves is used to detect and identify targets and objects of interest as well as locate, orient, and characterize them. This information can then be used to initiate computer applications and determine input data.

The method includes the step of using a distance equation (described in the definitions section) in this Color Ratio Space which is used to determine the presence of any of the N factorial Rich Color types that are used to identify the object or target that is being searched for. For a pixel located I, j the color distance equation is defined as follows in terms of ratio space color component values (c_1, c_2, \dots, c_N):

$$A_m = POS(B_{1,m} * (f(c_1) - T_{1,m})) + POS(B_{2,m} * (f(c_2) - T_{2,m})) \dots + POS(B_{N,m} * (f(c_N) - T_{N,m}))$$

Where, for color component n and color filter type m, $T_{n,m}$ is a color threshold, $B_{n,m}$ is the Bias, $f(c_n)$ is a function of c_n

35

and the function POS(A) such that POS(A)=0 if A is less than or equal to 0, ELSE POS(A)=A. Since only a binary result (zero or non-zero) answer is of interest, the POS functions above can be replaced with traditional IF tests.

Alternatively the method further includes the step of creating a corresponding look-up-table for each primary color and secondary color used in a target capturing an image and subtracting from each pixel in the image the bias of each camera color component apply the ratio space look-up-table to each pixel in the image for each primary and each secondary color used in the target to implement a color filter in this color ratio space to determine the presence of any of the N factorial Rich Color types that are used to identify the object or target that is being searched for.

A robust method using a computer or embedded computer chip or specialized camera circuitry to rapidly find the location of a target within an image by searching for Rich Color patches adjacent to each other within a tolerance.

This method can be implemented on a variety of platforms. It could be developed as a software program running on common personal computers, tablets or smart phones. It could be implemented at a hardware level as a dedicated chip that could be designed-in to hardware systems such as digital cameras. It could be implemented as a stand-alone appliance that could be retro-fitted to existing systems.

The method described herein can be used iteratively on the same data to find the optimal thresholds and bias parameters.

This method can be extended to include more than just the typical three sensor channels (R,G,B). It is likely that digital cameras will soon be commercially available with four sensor channels and cameras with even more channels will be available later in the decade. Furthermore, Security cameras commonly include infrared sensors.

Poor quality data streams (for instance those shot in very low light) could be enhanced to emphasize their rich color characteristics prior to processing by this method in order to allow the processing of data that might otherwise be unusable.

The method further includes the step of setting an indicator for each pixel that identifies the type of Rich Target Color present or the lack of a Rich Target Color.

The method further includes the step of identifying a pixel as having a non-Rich Color if each of its input color components is less than a black tolerance.

The method further includes the step of identifying a pixel as having a non-Rich Color if each of its input color components is above a white tolerance.

The method further includes the step of filtering image data by limiting image data to patches of Target Color that flank each side of a Rich Color Transition Point.

A smart camera with this method implemented in hardware could analyze each frame while the next frame is being acquired and then transmit (wired or wireless) a small data set of Transition Curves only if a target is detected in the frame. This method requires only a small amount of buffer memory and a very few multiples, adds, if tests and bit manipulation per pixel processed. This leads to a blindingly fast method for finding targeted objects in an image. The method's speed and flexibility make it possible to use a variety of targets at the same time. If targets with unique color and geometry are used, then each target that is found can trigger a unique action based on their absolute or relative location and orientation together with their associated database properties.

What is claimed:

1. A threshold color transition curve tracking apparatus for use in a method identifying a target and an image using a

36

transition curve formed of threshold color transition points between two color types on the target comprising:

a carrier presented on a surface; and

a plurality of unique threshold color transition targets arranged in a unique sequence on the carrier, the targets allowing identification of the location and orientation of the surface on which the targets are presented, by a camera.

2. The apparatus of claim 1 wherein: the carrier is an elongated strip.

3. The apparatus of claim 1 wherein:

the targets vary from each other in at least one of shape, color, patterns, area, angle of orientation on the carrier and separation distance from adjacent targets on the carrier.

4. The apparatus of claim 1 wherein:

the targets have one of a polygonal shape and a concentric circle shape.

5. The apparatus of claim 1 wherein:

the targets are aligned along a centroid of the carrier.

6. The apparatus of claim 1 wherein:

the targets comprise a first series of targets aligned along a centroid of the carrier; and satellite targets spaced from the centroid of the carrier.

7. The apparatus of claim 1 wherein the targets comprise: a multi-dimensional array of targets on the carrier.

8. The apparatus of claim 1 wherein the carrier comprises: a variable electronic display.

9. The apparatus of claim 1 further comprising:

location markers presented on the carrier at incremental dimensions along the carrier.

10. The apparatus of claim 1 wherein:

the colors on the targets are rich colors; and

the threshold color transition curves are rich color transition curves.

11. A method for identifying the location and orientation of an object in an image using a threshold color transition curve tracking apparatus to identify a target in an image using a transition curve formed of threshold color transition points between two color types on the target comprising:

providing a plurality of unique threshold color transition targets arranged in a unique sequence on a carrier;

attaching the carrier onto a surface;

identifying the surface on which the unique targets and sequence of targets using the unique targets and the sequence of targets to identify the location and orientation of the surface on which the set of targets are attached.

12. The method of claim 11 further comprising:

storing data defining the color transition curves corresponding to the unique set of targets arranged in the unique sequences in a computer accessible memory;

using a computer processor to compare a unique set of targets on one surface captured by the image capturing device with the set of data defining the transition curves of the stored unique set of targets to identify the target and the location and orientation of the target captured in the image.

13. The method of claim 11 further comprising:

Forming the carrier as an elongated strip.

14. The method of claim 11 further comprising:

varying the targets from each other in at least one of shape, color, patterns, area, angle of orientation on the carrier and separation distance from adjacent targets on the carrier.

15. The apparatus of claim **3** further comprising:
forming the targets with one of a polygonal shape and a
concentric circle shape.

16. The apparatus of claim **3** further comprising:
aligning the targets along a centroid of the carrier. 5

17. The apparatus of claim **3** further comprising:
providing the targets as a first series of targets aligned along
a centroid of the carrier; and
providing satellite targets spaced from the centroid of the
carrier. 10

18. The apparatus of claim **3** further comprising:
forming a multi-dimensional array of targets on the carrier.

19. The apparatus of claim **3** further comprising:
forming the carrier as a variable electronic display.

20. The apparatus of claim **3** further comprising: 15
providing location markers presented on the carrier at
incremental dimensions along the carrier.

* * * * *